Constrained Own-Account Work in Developing Markets: Evidence from Brazil*

Thiago Scarelli†
Paris School of Economics (PSE)

Abstract

This study explores a fundamental trade-off imposed by imperfect labor markets: individuals may work on their own at any time, but they can only occupy a potentially more productive wage job after a search period. We formalize this intuition using a simple extension of the canonical job search framework, which leads to a set of implications that can help to explain the prevalence of own-account work in developing labor markets. In particular, we show that a sufficiently high time discount rate can rationalize the puzzling choice of own-account work when it offers a lower instantaneous return relative to wage employment. In the second half of this research, we use this theoretical structure to empirically characterize the minimum discount rate that is consistent with the observed occupational decisions of Brazilian own-account workers given the wage employment opportunities they potentially face. In our baseline specification, we find that in nearly 70% of the cases the lower bound implied by the observed choices is strictly above the rates available on the credit market, which we interpret as evidence of a financially constrained occupational choice. This result suggests that the majority of own-account work in Brazil is driven by the combination of labor market frictions and financial market failures.

JEL Classification: J22, J24, J31, J64.

Keywords: Self-Employment, Own-Account Work, Imperfect Labor Markets, Job Search, Developing Countries, Financial Constraints, Time Discounting.

---

*I am grateful to David Margolis for his guidance and support throughout this research. I also thank François Fontaine, Sylvie Lambert, Luc Behaghel and Thomas Breda for their helpful comments and discussion. Remaining errors are my own.

†Email: thiago.scarelli@psemail.eu

1 Introduction

The International Labor Organization (ILO) estimates that about one-third of the employed population worldwide is composed of own-account workers. Despite its magnitude, this particular form of employment is largely underrepresented in labor economics research, being often subsumed under other topics, be it "entrepreneurship" or "informality". As a consequence, we still lack a clear view of the distinctive features of own-account work and why is it a much more prevalent form of labor supply in poor regions. The objective of the present study is to advance our understanding on the determinants of own-account work in the context of developing labor markets, exploring the link between occupational choice, labor market frictions and financial markets constraints.

Our contribution is twofold. First, we formalize the agents' decision to work on their own using an extension of the canonical job search model in partial equilibrium. The key innovation here is to frame autonomous work as an alternative that is immediately available to the individual, so that the occupational decision boils down to the comparison between the discounted value of own-account work, given the agents' own productivity, and the discounted value of unemployment, given the conditions available in the wage employment market. This simple model offers some realistic implications that cast light on how the share of own-account workers in the workforce may be driven by a parsimonious set of labor market parameters.

Second, we use the empirical counterpart of this model to infer how many own-account workers in Brazil can be said to be financially constrained. Our strategy is to estimate separately the expected value of each component of the model and then to use the revealed choice of the own-account workers to structurally identify the lowest time discount rate that is compatible with their decision. The intuition is that agents with a low time discount rate may prefer to wait a few periods for a well-paying wage job, while those with a very high preference for present consumption would be more willing to bypass the search process and start working on their own right away. Therefore, by observing the reported income of own-account workers and comparing it to their potential earnings in the wage jobs held by people with similar observable attributes, we can infer how high their rates of time discounting must be. Using this strategy, we find that nearly 70% of the Brazilian own-account workers have a time discount rate that is strictly superior to the market interest rates, which we interpret as evidence that their occupational choice was driven by a financial constraint.

This discussion is relevant for policy making because it suggests that scarce job opportunities combined with imperfect credit markets can contribute to the high prevalence of autonomous work in poor regions, challenging the traditional narrative according to which workers simply self-select into the occupation that offers the highest instantaneous return. If today's subsistence needs take precedence over the possibility of finding a well-paid job tomorrow and intertemporal exchange of consumption is not available, rational agents can be persistently trapped in unproductive tasks. According to our estimates, this is not a marginal possibility — it can be the main driver for the majority of own-account workers in a developing country.

Our study also suggests that there is wide heterogeneity in the time discount rate used by individuals to make occupational decisions. This is relevant because there is scarce empirical evidence on the distribution of subjective discount rates, notably so in the developing world, and applied labor research often assumes one homogenous rate for all agents. Our findings suggest that this simplification can be misleading, especially for workers out of urban centers, for whom the relevant discount rate can be very distant from the market rate.

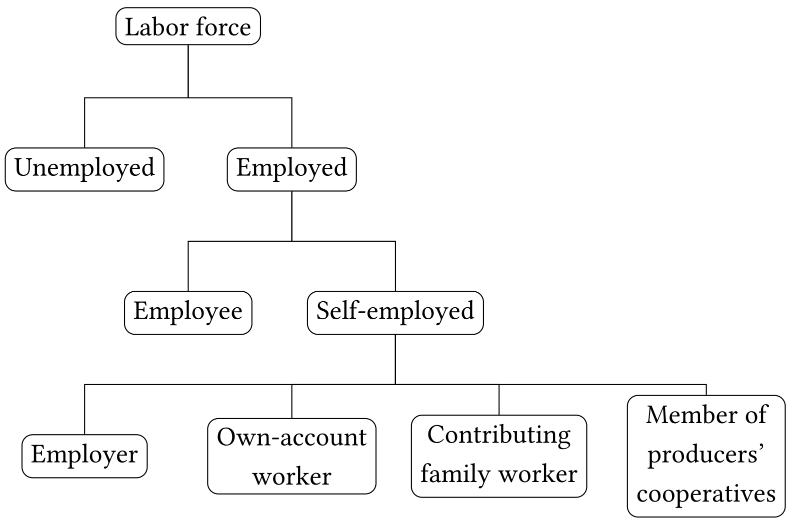
We highlight that our empirical results are valid conditional on two fundamental assumptions: (a) the model captures the relevant components of the occupational decision and (b) on average, our estimates of the labor market parameters are sufficiently similar to how workers perceive their own wage job opportunities. In a nutshell, we assume our expected values (in the statistical sense) to be a translation of the values expected by the individual (in a heuristic sense). To address robustness concerns, we present in detail how the results change with different specifications, stressing that our estimates are particularly sensitive to the way we estimate the potential wage offers.

The remainder of the study is organized as follows. Section 2 examines a precise concept of own-account work as a subcategory of self-employment and presents some descriptive statistics on the prevalence of this occupation both across countries and in the particular case of Brazil. Section 3 reviews three related branches of the labor literature: (a) the neoclassical "entrepreneur" models and why they are ill-suited to explain the specificities of own-account work, (b) the debate over "necessity" versus "opportunity" self-employment, with a focus on the developing world, and (c) how self-employment is usually interpreted in the job search framework. Section 4 presents the Brazilian labor market survey that we use in the empirical estimation. In the core of this study, Section 5 introduces the theoretical model, discusses its implications and presents the empirical estimation strategy. The results of each regression and our findings on the distribution of discount rates are examined in Section 6, with the variations on the baseline specification being examined in section 7. Finally, section 8 discusses the limitations of the study and concludes.

2 Concepts and Stylized Facts

The core subject of this research is *own-account work*, which can be briefly defined as a self-employed without employees.[1] This concept has two important components. First, it states that own-account work is under the more general category of *self-employment*, understood as the case where the individual has authority over the organization of her economic activity and the remuneration is directly dependent upon its profits. This contrasts to the case of *paid employment*, whose basic remuneration is defined by a contract with a third party, which is the one who bears the largest share of the risks and is also entitled to the returns of the activity. Second, the concept emphasizes the strict self-reliance of the own-account worker, in contrast to the *employer* (who engages other employees on continuous basis), the *contributing family worker* (a young relative, typically working without pay) and the *members of producers' cooperatives* (who share responsibilities and returns among themselves in equal conditions), as summarized in figure 1. In short, own-account workers are characterized by a double autonomy: they are autonomous to organize their economic activity according to their will, but also in the sense that they are the sole agent engaged in it.

Figure 1: *Schematic overview of the different self-employment categories.*

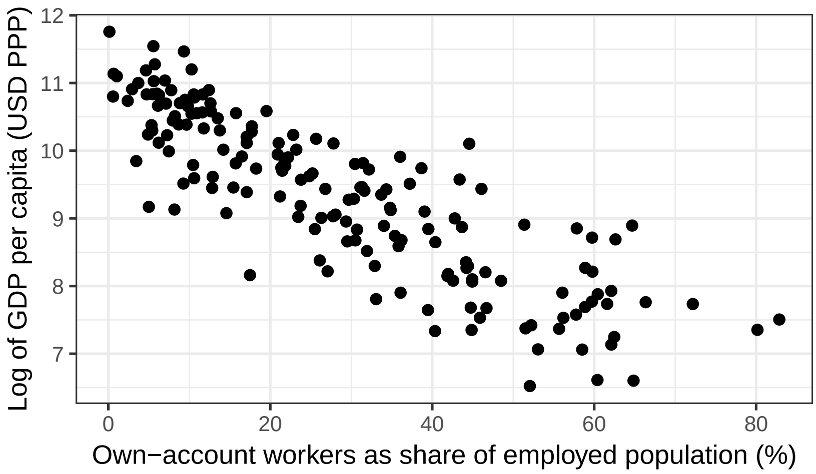


[1] The International Classification of Status in Employment (ICSE-93), which is the current reference for the ILO, states that "[o]wn-account workers are those workers who, working on their own account or with one or more partners, hold the type of job defined as a 'self-employment job' (...), and have not engaged on a continuous basis any 'employees' (...) to work for them during the reference period. It should be noted that during the reference period the members of this group may have engaged 'employees', provided that this is on a non-continuous basis." (Hoffmann, 2003).

We insist on this distinction among the categories of self-employment because its empirical content is often overlooked. Even if we put aside the contributing family workers (who are typically not remunerated) and the members of cooperatives (which generally comprise a negligible share of the employed population), there still remains two groups under self-employed which in practice are very distinct. As we are going to see later in section 4, own-account workers in the developing context frequently have a weak human capital, are often involved in non-specialized tasks and dominate the lower end of the income distribution, while the opposite is true for employers. If those subcategories have indeed distinct dynamics, there is a methodological reason to investigate own-account work as a subject of its own, as we do in this research.

The first empirical regularity that emerges when we look at the presence of own-account workers across different labor markets is that their prevalence is inversely proportional to the average income in that market. For concreteness, recent ILO estimates show that the majority of the employed population in poor countries are own-account workers (e.g. 83% in Haiti, 60% in Uganda and 54% in Somalia), in comparison to small minority in the high-income group (e.g. 7.1% in France, 4.6% in Norway and 3.4% in the United States).[2] To be sure, there is Pearson correlation of -0.86 between the share of own-account workers and the log of the GDP per capita in USD PPP terms, as illustrated in figure 2.

Figure 2: *Share of own-account work in the employed population and GDP per capita, per country, as of 2017. Sources: Occupation data comes from the ILO; income data comes from the International Monetary Fund (IMF) World Economic Outlook (WEO) database.*



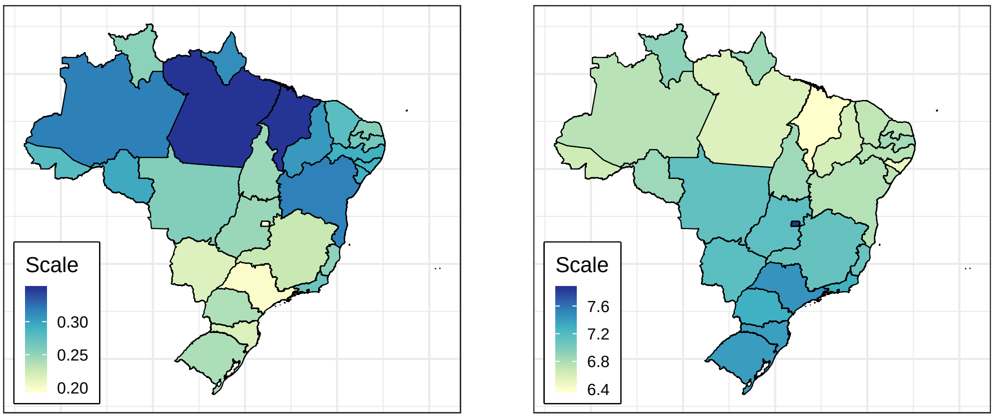
In this respect, Brazil is an interesting case because, as a whole, the country has a 25% share of own-account workers, not far from the global average, and its

[2] The estimates are available at http://www.ilo.org/ilostat-files/Documents/Excel/MBI_32_EN.xlsx.

regions reflect the same pattern observed between countries. Overall, the States in the south have a relatively higher income and a lower share of own-account workers, as shown in figure 3. In the Federal District and in the State of São Paulo, two of the richest areas, own-account work responds for 20% of the employed population, a level somewhat comparable to the South Korean average, while in the least developed areas, as the hinterland of Maranhão, Amapá or Amazonas, this proportion nearly doubles, approaching the levels estimated for Rwanda.

Figure 3: *Share of own-account work and average income in Brazil, per State, as of 2017. Source: Own calculation using PNAD microdata.*

(a) *Share of own-account workers* **(b)** *Log of average monthly income per capita*



This clear pattern helps us to put in context the discussion about the autonomous supply of labor. From the developed world perspective, it is tempting to picture the workers who are outside the logic of the usual labor contract as managers of dynamic start-up companies, while in poor countries this employment category actually comprises a wide range of workers, many of them close to the subsistence level. Second, this stylized fact leads to the question of why is this so and what characteristics of those markets may contribute to such a strong presence of own-account workers. Finally, it points to the importance of assessing to what extent those workers can be said to be in their first best occupation and what kind of constraints they could be facing. In the following section, we review the explanations offered by the literature, before examining a new formalization of this issue using the tools of the job search framework.

3 Literature Review

The research on own-account work as an independent topic is still remarkably underdeveloped and unstructured. As an illustrative example, the second edition of the textbook “Labor Economics” (Cahuc, Carcillo, and Zylberberg, 2014), which is a reference for the field, does not mention this term a single time over its 1,043 pages. Even the broader idea of self-employment is largely absent: the terms “self-employment” or "self-employed” are used only 8 times in the book, mostly as an auxiliary concept in the definition of unemployment.

For that reason, our work does not fit within a single branch of the literature but instead relates to three different debates. In what follows, we briefly go over the classic theories that see self-employment as entrepreneurship, in order to argue why they cannot fully explain the phenomena of own-account work in the developing context. Then we move on to consider some institutional determinants of own-account work and the “opportunity” versus “necessity” debate. Finally, we comment on how self-employment usually enters in the basic job search model.

3.1 Self-employed as a skilled, risk-lover and liquid entrepreneur

Our research is related to the entrepreneurship literature by the common interest in understanding why some agents take the role of employees and sell their workforce, while others have the initiative to bring goods and services to the market on their own.

The most traditional explanation, which goes back to Roy (1951), emphasizes the role of skills and relative productivity. In short, self-employed are those for whom the return to entrepreneurship is relatively higher than the return to paid work. Along those lines, Lucas (1978) suggests that a random distribution of managerial “talent” over the population would be sufficient to determine the division of labor into managers and employees, as well as the size of the firms under each manager. As a possible refinement, Lazear (2005) proposed that entrepreneurs are actually not endowed with a particular skill, but instead are competent in many: self-employed are “jacks-of-all-trades” and employees are experts.

In complement to skills, there is also long literature pointing to the importance of individual attitudes towards risk. Formalizing the perspective from Knight (1921) on the role of the entrepreneur, Kihlstrom and Laffont (1979) show that if individuals must choose between a fixed wage and a risky profit, self-employed and employees will be sorted by their degree of risk aversion. A set of empirical papers have been able to establish an association between self-employment and psychological measures of tolerance to risk (see Evans and Leighton, 1989; Ekelund et al., 2005; Masclet et al., 2009), but the direction of causality is unclear, as the experience with autonomous work in itself can change how one manages risk.

One could expect that, in the presence of such high risks, entrepreneurs would require a premium, as suggested by Kanbur (1982). Yet, the empirical evidence suggests the opposite is true: they bear more risk but have a lower average income and a lower income growth relative to wage workers. Using data from the United Stated, Hamilton (2000) estimates a median earnings differential of 35 percent for individuals in business for 10 years, without accounting for employee fringe benefits such as employer-provided pension or health insurance.

One way to justify the preference for self-employment despite its lower returns is to consider that individuals may derive satisfaction from non-monetary aspects of this occupation, such as the flexibility, the autonomy and the sense of control that comes with "being your own boss". Indeed, self-employed individuals in OECD countries tend to report higher life satisfaction, even if they complain of a heavier workload (Blanchflower, 2004), suggesting that they could be willing to exchange monetary compensation and leisure time for the possibility to have control over their economic activity. In any case, it is still not clear what is the direction of the causality: self-employed can be more satisfied, but happy people can be more likely to prefer self-employment.

Finally, there is a body of literature suggesting that, in the absence of well-functioning capital markets, the decision to move into self-employment becomes contingent on the availability of personal funds. One classic reference in this literature is Evans and Jovanovic (1989), who suggested that, in the context of the United States, liquidity constraints prevent skilled self-employed from starting a new business and limit the size of the firm for those who try it. Since then, this particular hypothesis has found support from reduced form estimations exploring exogenous access to liquidity: Blanchflower and Oswald (1998), using data from the United Kingdom, find that individuals who receive an inheritance are more likely to become self-employed, and Lindh and Ohlsson (1996) find a similar result for lottery-winners in Sweden. In both cases, if the liquidity aspect were not binding, windfall gains would have no effect.

In summary, the discussion on entrepreneurship suggests that self-employment is most often taken by individuals who are relatively more skilled in it, are tolerant to risk, have sufficient access to liquidity and are willing to get a lower monetary return in exchange for autonomy. As a whole, this framework appears to be adequate to think about the dynamics of new business creation in developed markets — to be sure, the reference papers we mentioned so far in this section are built upon empirical evidence from the US, the UK, Germany, and Sweden.

On the other hand, it is not clear to what extent the entrepreneurship view is able to explain the prevalence of own-account work in poor regions. Can we claim that Rwanda, Honduras and the northern region of Brazil have a large share of own-account workers because they offer a particularly efficient capital market? Or because the local population has a disproportional endowment of managerial talent? Autonomy and flexibility can be important factors if the material conditions of subsistence are secured, but probably less so for own-account workers living with less than the minimum wage, as we observe in Brazil. Hence, to the extent that developing markets have peculiar dynamics, it is worth examining alternative channels and complementary explanations.

3.2 Self-employment in the developing world

This present research is most closely related to a growing literature that stresses the major role played by own-account work in developing countries and the need to understand its particularities under the light of the local economic and institutional conditions.[3]

Part of this literature has used reduced form estimations to documented that self-employment is indeed a very heterogeneous category in poor countries, with well-educated employers and low-qualified own-account workers coexisting in the same economy. This is the case, for example, in Colombia (Mondragón-Vélez and Peña, 2010), in Argentina (Mandelman and Montes-Rojas, 2009), and in the transition economies in Eastern Europe (Earle and Sakova, 2000).[4] This branch of descriptive studies also suggests that a share of the own-account workers could be "disguised unemployed", but they do not formalize the trade-offs between the different work states, and they do not ask what exactly would be preventing them from pursuing a different occupation.

Taking a broader view, Margolis (2014) argues that the incidence and the type of self-employment observed in a given economy are mainly determined by four groups of factors:

---

[3] The most representative studies in this small literature are Earle and Sakova (2000), Mandelman and Montes-Rojas (2009), Mondragón-Vélez and Peña (2010), Margolis, Navarro, and Robalino (2012), Grimm, Knorringa, and Lay (2012), Monsen, Mahagaonkar, and Dienes (2012), Narita (2013), Margolis (2014), Gindling and Newhouse (2014), and Cho, Robalino, and Romero (2015). In many aspects, this discussion can also be seen as part of a broader debate on the particularities of the labor markets in poor regions, as presented in Fields (2011) and Banerjee and Duflo (2011), particularly their discussion on the overlap between self-employment, informality, and entrepreneurship in this context.

[4] The same pattern holds also in the Brazilian case, as we will present in section 4.

1. the presence of *social protection mechanisms*, either formal or informal, which can ensure minimal levels of consumption and prevent individuals from starting precarious self-employment to avoid starvation;

2. the extension of *labor market frictions*, which make access to information about jobs more costly and may increase the effective monopsony power of employers;

3. the local *business environment*, including the costs of doing business, the access to capital, the tax policy and the effectiveness of the legal system, which affects the creation of jobs;

4. and the *labor market regulations*, such as the minimum wage, safety standards, payroll taxes, and employment subsidies.

The determinants of self-employment here go beyond individual skills and preferences and complement those aspects with considerations on the structure of the labor market. The articulation of those elements in formal models and the empirical measurement of their effects are currently at the forefront of this research agenda.

3.3 Measuring the "necessity" self-employment

Another concern in this literature is how to distinguish between "opportunity" and "necessity" self-employment — or, similarly, between "subsistence" and "transformational" entrepreneurship (Schoar, 2010), "pull" and "push", or "choice" and "constraint". This duality appears to be intuitive and yet it is very elusive to be defined in precise terms.

This difficulty is translated in the lack of consensus on how to measure it. At first sight, one could simply to ask people what is the main motivation behind their business initiative, which is the approach adopted by the Global Entrepreneurship Monitor survey (see Poschke, 2013, for a cross-country exploratory analysis of this data source). In its most recent release for Brazil, referent to 2016, the survey reports that 42% of the early-stage entrepreneurs in Brazil have said to "have no better choices for work" (Global Entrepreneurship Monitor, 2017).[5]

[5] We note that the representativeness of this measure is limited by the usual bias related to self-reported measures of success, as individuals may feel induced to fulfill the expectations of the interviewer. Moreover, the sampling design included only 2,000 people in 27 cities, which does not provide much confidence that the results incorporate proportionally the conditions found in the rural areas of the country, where the labor market is less dynamic.

A second strategy is to define an objective measure of success and then classify self-employed workers according to it, assuming that those who are doing well (or who have the potential to do well) can be said to be opportunity-driven. To that end, Gindling and Newhouse (2014) explores two different criteria: whether the self-employed hires additional employees or has a family per capita income above the $2-per-day poverty line. Having identified those two reference groups, they estimate how many among the "unsuccessful" self-employed are sufficiently similar to the "successful" ones to be considered as "high-potential" cases. They find that about 64% of the non-agriculture self-employed worldwide are "low-potential", regardless of the criterium chosen. In particular, for the Latin America region, they suggest that low-potential self-employment could range from 53% to 60% of the cases, depending on the selected measure.

Finally, a third strategy is to look at the employment state of the individual prior to self-employment and classify as "opportunity" those who were previously employed, and as "necessity" those who were unemployed.[6] While appealing in its simplicity, it is not clear what is the empirical component and the magnitude of the measurement error of this strategy. Moreover, necessity self-employment becomes mechanically counter-cyclical due to the definition itself, which limits how much new information this approach can actually offer.

As we can see, there is yet no agreement on how to assign a precise definition to the opportunity-or-necessity duality that could lead to a consensual measurement strategy. The present work aims to offer a new contribution to this open question.

3.4 The job search framework

In terms of the theoretical methodology, the model we will discuss is built upon the basic job search model that followed the pioneering work of Stigler (1961) on the economics of search. In opposition to Marshallian models, this literature acknowledges that sellers and buyers do not have perfect information about each other, and that prices in the market are dispersed. Therefore, agents must go through a costly process of acquisition of information in order to trade.

This idea is particularly powerful when applied to the labor market, where job-seekers need to find a firm in order to trade their labor services, and it has led to a large search-theoretic literature following the works of McCall (1970) and Gronau (1971). This research agenda later evolved to an equilibrium theory, where

[6] Small variations of this rule are applied in the context of developed countries by Block and Sandner (2009) and Baptista, Karaöz, and Mendonça (2014), among others. In the most recent example, Fairlie and Fossen (2018) estimate that 80-90% of the entrepreneurs in the United States and in Germany are opportunity-driven.

job-seekers and firms find each other according to a matching mechanism, as developed by Diamond (1982), Mortensen (1982) and Pissarides (1984). In any case, as noted by Rogerson, Shimer, and Wright (2005), even the rudimentary versions of the models in this tradition are able to explain two facts that are absent in frictionless models: it takes time to find a job and there is a stochastic component in the workers' allocation in the market.

Even though we will ultimately stick to partial equilibrium in our model, it is instructive to briefly review how own-account work is traditionally disregarded in the broad search-and-matching framework. Mortensen (1987) goes as far as considering that the unemployed are liquidity constrained and self-finance their search period, but assumes that "the worker's only alternative when the liquidity constraint is binding is to drop out of the labor force" (p. 860). In the standard version of the model, own-account work implicitly has an accessory role: Pissarides (2000) defines the instantaneous income received while in unemployment (z, in their terminology) as including the unemployment insurance and the earnings from "odd and irregular jobs in a secondary sector of the economy, if such sector exists" (p. 13). Hence, by framing it as a temporary task one might have *while searching*, there is no room for persistent own-account work and there is no trade-off between this employment category and wage employment. In any case, the only alternative to job search is assumed to be non-participation — and this is precisely the gap we aim to explore.

Motivated by a similar objective, Albrecht, Navarro, and Vroman (2009) added informality to the model from Mortensen and Pissarides (1994). Effectively, they extend the framework to include "unregulated self-employment", in a sense that could be understood as own-account work. However, their modeling choice differs from our approach in two important ways. First, they assume that the income flow at the informal sector is the same for all workers, while we take the opposite stance and model the return at own-account work as individual-specific productivity. Second, unemployed workers can match with a firm or can take up an informal opportunity, *if they find one.*[7] In practice, this strategy amounts to modeling sectors with distinct arrival rates and distinct instantaneous returns, but that otherwise at not essentially different. Our approach, in contrast, highlights that the fundamental feature of own-account work is the possibility to bypass the search period, stressing its role as an outside option to unemployment.

[7] A variation of this approach is adopted by Margolis, Navarro, and Robalino (2012) to analyze the Malaysian labor market. In their model, job-seekers receive offers from both the formal and the informal wage job sectors, and can also find self-employment opportunities, at a given rate.

4 Data

Our empirical analysis is based on the public microdata from Brazil's "Pesquisa Nacional por Amostra de Domicílios" (PNAD), a nationally representative labor market survey that collects information about nearly 560,000 individuals per quarter.[8] This source is appropriate for our research interest because it is one of the few surveys in the developing world to offer a relatively large sample size with a longitudinal structure and retrospective information on the duration of individual's employment status.

We focus on the 8 quarters of data from 2016 and 2017, which originally includes 4,53 million observations of 1,6 million individuals. After extracting the socioeconomic data from all interviews, we remove individuals with less than 15 years of age (from whom labor information is not collected) and those missing data on race or education. We also trim the observations that are extreme outliers with respect to work income, dropping the top and the bottom 0.1%. The final sample covers 3,56 million observations from 1,26 million individuals.

The design of the survey establishes that each housing unit is to be visited for five consecutive quarters before leaving the sample. The public microdata files already include a unique identifier for the household level that is common across interviews. We take a step further and match the individuals within a given household using their gender and birthdate.[9] Hence, our unique individual identifier allows us to follow the individuals for up to 15 months and to identify any changes in their labor market status within that window.

Table 1 presents the summary statistics for the working sample as a whole and also by group according to the individual labor market status, as this segmentation allows us to have a better idea of the different profiles in each category.

[8] For a detailed description of the survey, see IBGE (2014, 2016, 2018).

[9] Using gender and strictly identical birthdate within a household, we can assign unique identifiers and match about 93% of the observations. For the remaining cases, which have no birthdate available, we infer a plausible date from their reported age and then try to match them with the previously identified individuals with the same gender and similar age. Finally, we try to match the remaining unidentified individuals among themselves, again using the same gender and similar age criteria. To be precise, we allow the acceptable gap in "similar age" to be increasing according to the rule $(age/25)^2 + 1$, in the spirit of Ribas and Soares (2008). This rule basically says that we tolerate about one year of difference to match people around 25 years old, but up to 5 years gap to match people at the age of 50, to reflect the accuracy loss in the imputation of assumed age for older individuals. After this process, we have that 27% of the individuals are observed a single time, 21% appear 2 times, 17% appear 3 times, 15% appear 4 times, and 21% appear in all 5 interview rounds.

Table 1: *Summary statistics by work state*

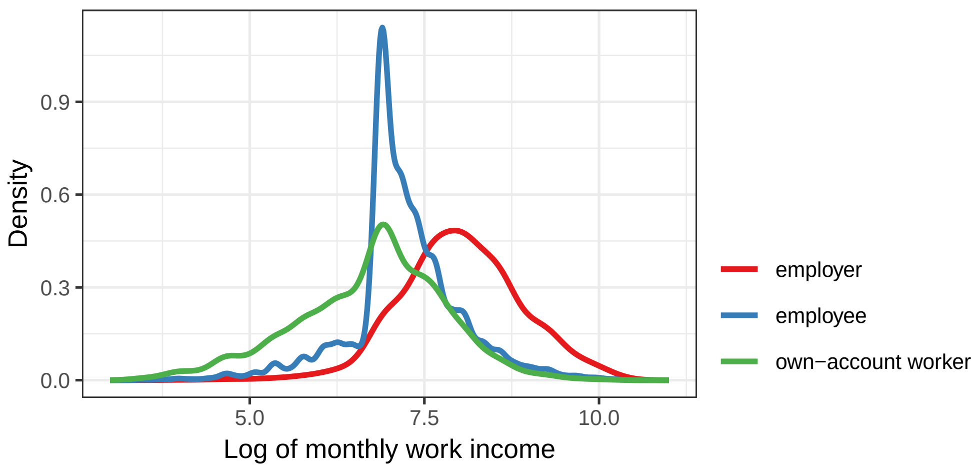
	Full sample	Employee	Own-account worker	Employer	Unempl.	Inactive	Family worker
Age (in years)							
Age (mean)	41.82	37.28	43.86	45.89	30.70	46.98	35.10
Age (median)	40.00	36.00	44.00	45.00	28.00	49.00	32.00
Age (std. dev.)	18.03	12.32	13.60	12.64	12.06	22.34	15.92
Gender (in %)							
Male	0.48	0.55	0.70	0.71	0.50	0.34	0.36
Female	0.52	0.45	0.30	0.29	0.50	0.66	0.64
Race (in %)							
White	0.41	0.42	0.40	0.63	0.32	0.39	0.40
Non-white	0.59	0.58	0.60	0.37	0.68	0.61	0.60
Highest educational level (in %)							
No school	0.11	0.04	0.11	0.03	0.04	0.18	0.08
Elem. school incomplete	0.31	0.22	0.39	0.19	0.23	0.38	0.44
Elem. school	0.10	0.09	0.10	0.08	0.10	0.10	0.12
High school incomplete	0.08	0.07	0.06	0.04	0.12	0.09	0.11
High school	0.25	0.33	0.24	0.31	0.36	0.17	0.20
Some college, no degree	0.04	0.06	0.03	0.06	0.06	0.03	0.02
Graduate or above	0.11	0.19	0.08	0.29	0.08	0.05	0.03
Literacy (in %)							
Literate	0.91	0.97	0.91	0.98	0.96	0.84	0.92
Illiterate	0.09	0.03	0.09	0.02	0.04	0.16	0.08
Currently going to school (in %)							
Not at school	0.87	0.89	0.96	0.96	0.81	0.82	0.81
At a public school	0.09	0.05	0.02	0.01	0.13	0.15	0.17
At a private school	0.04	0.06	0.02	0.03	0.06	0.04	0.02
Urban status (in %)							
Live in urban area	0.73	0.82	0.65	0.87	0.83	0.69	0.27
Live in rural area	0.27	0.18	0.35	0.13	0.17	0.31	0.73
Earnings (in BRL)							
Work income (mean)	1,844	1,882	1,400	4,321	-	-	-
Work income (median)	1,196	1,259	967	3,132	-	-	-
Work income (std. dev)	2,269	2,204	1,731	4,143	-	-	-
Sample size (in units)							
Number of observations	3,560,263	1,220,290	528,667	76,272	242,476	1,421,525	71,033
Number of individuals	1,259,125	527,648	264,063	40,201	162,765	598,712	43,506

Source: Own calculation using PNAD microdata from 2016 and 2017. Notes: The sample excludes individuals below 15 years old and trims work income outliers, dropping the top and the bottom 0.1%. All monetary values are adjusted to the equivalent purchasing power of December 2018. The race category “white” combines the responses “white” and “asian”, while “non-white” combines “black”, “brown” and “indian”. Since individuals can change state over the interview rounds, they may appear in more than one group and hence the sum of unique individuals for each group is superior to the count of unique individuals for the full sample.

In line with the usual finding in the self-employment literature (Blanchflower, 2004; Terrell and Troilo, 2010), both the own-account workers and the employers in Brazil are predominantly male (about 70%) and are generally older than the employees. Apart from those commonalities, their other attributes are very distinct. The group of own-account workers has a racial composition and schooling profile that is broadly comparable to the general population, but they are overrepresented in rural areas and report a median work income that is 20% below the unconditional median. On the other extremum, employers are much more likely to be white (63% versus 41% in the full sample), have a higher educational level (29% of graduates versus 11%), are concentrated in urban centers (87% versus 73%), and report a median income that is 2.6 times the national level.

An examination of the full distribution of work income by employment category provides more details on those earnings patterns. Figure 4 shows a mode at the minimum wage (about 1,000 BRL, or 225 EUR) for the employees, suggesting that this regulation is a focal point of earnings in wage jobs. For own-account workers and employers, the distribution of earnings is much more symmetric (but with a very distinct center of mass) and are less guided by the minimum wage, since the returns for those occupations are not contractually defined.

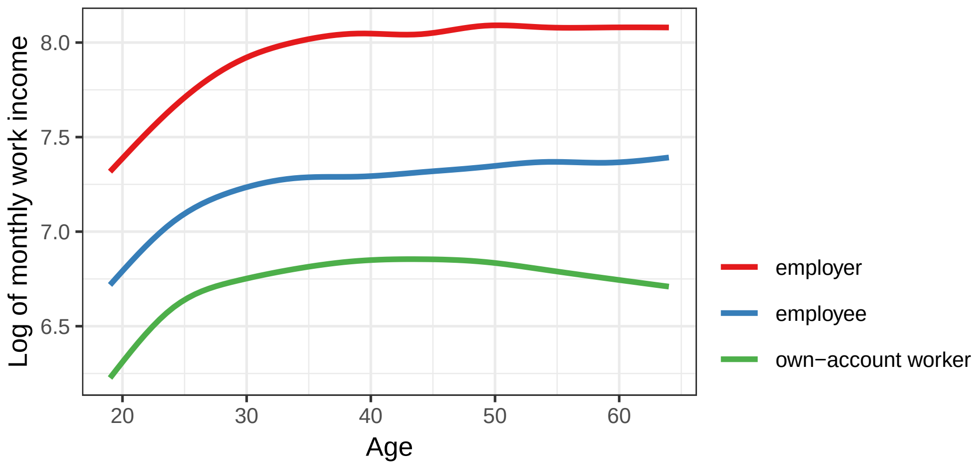
Figure 4: *Income distribution for employers, employees and own-account workers. Source: Own calculation using PNAD microdata.*



The distribution above also shows that there exist a sizeable share of own-account workers at the right tail of the earnings distribution who make as much as the employees or the employers. Those cases would fit well into the traditional explanation of self-selection driven by comparative advantages. The puzzle resides in the heavy left tail of own-account workers, which drives the average earnings gap relative to the other occupations. Such concentration of own-account workers at the low end of the income distribution provides suggestive evidence that this is a typical form of employment for low-return activities, and implies that factors beyond monthly earnings may influence the occupational choice of own-account workers, as we will discuss in the next section.

To conclude this section, we look at the average work income by age for those three employment categories. One could argue that the low income observed for a young own-account worker reflects the difficulties of a start-up business and that people would be willing to endure it in order to access higher returns later in life. This hypothesis finds little support in the data: the earnings gap observed around the 20s between own-account workers and employees remains relatively stable during the prime working age and spreads further after the age of 45. Even though this graph is not following a single cohort, we interpret these results as evidence that there are no extraordinary returns for experienced own-account workers, while the average work income of employees increases continuously, even if at a slow pace. This pattern may reflect an accumulation of occupation-specific skills but is also consistent with the experience-earnings profile required to provide career incentives to paid employees, as suggested by Lazear (1979).

Figure 5: *Average income by age for employers, employees and own-account workers. Source: Own calculation using PNAD microdata.*



All in all, this evidence supports our claim that own-account work is a particular employment category and constitutes a research object in itself. For that reason, in what follows we focus only on the relationship between own-account workers and wage employees, which are the natural reference point in the labor market. To be more specific, we will explore a theoretical mechanism that explains how it is possible that own-account work can be an attractive employment possibility even in the presence of better-paid wage employment, as we just described.

5 Theoretical Framework

The argument we formalize in the theoretical model is built around one fundamental intuition: individuals willing to supply their labor services must decide whether to work on their own or to look for a job position elsewhere. In the first case, people may enter their occupation immediately and will be entitled to all the income generated by their productivity. In the second case, the job-seeker need to find a firm willing to hire them and their income will be the wage agreed upon between both parties. Due to imperfect information in the labor market, the search process takes time.

We purposefully assume away any taste parameter — in our setup, the only determinant of utility is the monetary return of the occupation. From a methodological perspective, our challenge is to justify the choice for own-account work without relying on an ad hoc hypothesis about tastes, and without violating individual rationality.[10]

5.1 A model of own-account work within the job search framework

The model presented here is a simple extension of the job search framework in partial equilibrium.[11] As usual in the basic models of this literature, the environment is assumed to be stationary and the individuals are assumed to know the relevant labor market conditions. To be precise, they know the exogenous distribution of wages offered by the firms, how often one might get a job offer when looking for it, and how long the jobs usually last. The uncertainty lies in the fact that the actual characterization of each particular vacancy is unknown until the arrival of the job offer, which follows a stochastic Poisson process. Importantly, we add that individuals also know what is their deterministic productivity if they were to work on their own.

Agents have some degree of preference for the present, in the intuitive sense that $100 today is preferred to $100 in one month. For that reason, any future

[10] In any case, there might be no reason to assume that own-account work always offers a higher non-monetary utility than other forms of work. Hanglberger and Merz (2015) argue that the usual analysis overestimates the satisfaction of the self-employed because it ignores anticipation and adaption effects. The intuition is that people tend to be strongly dissatisfied before changing jobs and disproportionately content in the initial periods of the new one, a honeymoon phenomenon that could bias the comparison with wage jobs if the recent self-employed are overrepresented. Accounting for those dynamics, they estimate that the gap in satisfaction vanishes.

[11] A review of the broad class of search models can be found at Rogerson, Shimer, and Wright (2005). For a concise presentation of the elements we adopt in the present study, see chapter 5 of Cahuc, Carcillo, and Zylberberg (2014).

flow of income is discounted by a rate ρ that converts it into a comparable present value. Importantly, individuals are heterogeneous in their subjective discount rate.

One restrictive assumption we make is that individuals do not look for a job if they are working, which translates the idea that job searching requires an amount of effort and time that cannot be reconciled with the ongoing occupation. Furthermore, we abstract from the details of the matching mechanism and from any optimization behavior at the side of the firms. In this sense, we are adopting the idea of an optimal stopping rule, as in McCall (1970), where individuals sample from a given distribution of wage offers and stop searching whenever they find an offer above their reservation threshold.

5.1.1 Value of wage employment

The discounted value of any wage job $W(w)$ depends on the instantaneous wage it pays w and accounts for the possibility that the job may end with an instantaneous rate δ, in which case the worker would go back into unemployment, which has value U. Denoting a small time interval by dt, we can derive the usual flow value expression for employment as:

$$W(w) = \left(\frac{1}{1 + \rho \cdot dt} \right) \cdot \left[w \cdot dt + \delta \cdot dt \cdot U + \left(1 - \delta \cdot dt \right) \cdot W(w) \right] \tag{1}$$

$$W(w) + \rho \cdot dt \cdot W(w) = w \cdot dt + \delta \cdot dt \cdot U + W(w) - \delta \cdot dt \cdot W(w) \tag{2}$$

$$\rho \cdot W(w) = w + \delta \cdot \left[U - W(w) \right] \tag{3}$$

5.1.2 Value of unemployment

The discounted value of unemployment U (or, equivalently, the value of looking for a wage job) is given by the instantaneous unemployment benefit b the job-seeker may receive and by the expected gain from finding a job that pays w, given that at rate λ the unemployed draws an offer from the known distribution $F(w)$. Hence:

$$
\begin{aligned}
U = \left(\frac{1}{1+\rho\cdot dt}\right)\cdot\Bigg[& b\cdot dt \\
& + \lambda\cdot dt\cdot\int_0^{w_r} U\ dF(w) \\
& + \lambda\cdot dt\cdot\int_{w_r}^{\infty} W(w)\ dF(w) \\
& + \left(1-\lambda\cdot dt\right)\cdot U\Bigg]
\end{aligned}
\tag{4}
$$

Importantly, the equation above acknowledges that a job offer is only acceptable if it pays more than a given reservation wage w_r, defined as the lowest income necessary to make the individual indifferent between unemployment and wage employment. Therefore, any wage offer between 0 and w_r is refused and the individual remains unemployed, while offers above w_r lead to a job with a continuation value $W(w)$. With probability $1-\lambda\cdot dt$, the job-seeker receives no offer and continues to search.

Using the fact that $U = \int_0^{\infty} U\ dF(w)$, we have that:

$$
\begin{aligned}
\rho\cdot dt\cdot U = b\cdot dt & \\
& + \lambda\cdot dt\cdot\int_0^{w_r} U\ dF(w) \\
& + \lambda\cdot dt\cdot\int_{w_r}^{\infty} W(w)\ dF(w) \\
& - \lambda\cdot dt\cdot\int_0^{\infty} U\ dF(w)
\end{aligned}
\tag{5}
$$

Because $\int_0^{\infty} U\ dF(w) = \int_0^{w_r} U\ dF(w) + \int_{w_r}^{\infty} U\ dF(w)$, we can write the unemployment valuation simply as:

$$
\rho\cdot U = b + \lambda\cdot\int_{w_r}^{\infty}\left[W(w)-U\right]\ dF(w) \tag{6}
$$

5.1.3 The reservation wage

By definition, a job that pays the reservation wage has the same value as the unemployment state. Hence:

$$
\rho\cdot W(w_r) = \rho\cdot U \tag{7}
$$

Using equation (3) and the definition of the reservation wage:

$$w_r + \delta \cdot \left[U - W(w_r)\right] = \rho \cdot U \tag{8}$$

$$w_r = \rho \cdot U \tag{9}$$

Using the flow value of unemployment defined in equation (6):

$$w_r = b + \lambda \cdot \int_{w_r}^{\infty} \left[W(w) - U\right] \, dF(w) \tag{10}$$

Let us go back to equation (3) in order to rewrite $\left[W(w) - U\right]$:

$$\rho \cdot W(w) = w + \delta \cdot \left[U - W(w)\right] \tag{11}$$

$$\rho \cdot W(w) - \rho \cdot U = w + \delta \cdot \left[U - W(w)\right] - \rho \cdot U \tag{12}$$

$$\left[\rho + \delta\right] \cdot \left[W(w) - U\right] = w - \rho \cdot U \tag{13}$$

$$\left[W(w) - U\right] = \frac{w - w_r}{\rho + \delta} \tag{14}$$

Finally, substituting it back into equation (10):

$$w_r = b + \lambda \cdot \int_{w_r}^{\infty} \frac{w - w_r}{\rho + \delta} \, dF(w) \tag{15}$$

5.1.4 The value of own-account work

So far, the valuation equations follow the canonical results. In order to add the possibility of own-account work, we make three fundamental assumptions.

First, *own-account work is always available*, in the practical sense there is no need to wait for it. By definition, this is an autonomous decision that precludes coordination with third-parties. This assumption may seem strong, as one may argue that setting up a new activity may take time — for instance, it might be necessary to find clients. However, we note that someone looking for clients is *already occupied* doing so, is already an own-account worker, which is fundamentally different from a job-seeker waiting for a call-back.

Second, *the income generated by the own-account activity is given by the individual productivity*, which is an individual-specific parameter y that is sufficient to characterize the occupation. Because we assume all relevant utility from work is summarized by the monetary returns, it is possible for the individuals to rank all their possible alternatives (Should I paint houses? Offer private Math classes? Steal bicycles? Sell fruits in the street? Beg for money at the corner?) and y can be interpreted simply as the highest possible return in that list, given their idiosyncratic skills and the market constraints. Moreover, by definition, there are no principal-agent issues and no surplus to be shared, hence the worker is simply entitled to the full profit y.

Third, there is no exogenous destruction rate. To be precise, *the possibility of a destruction rate is immaterial for the valuation*, which is a logical consequence that follows from the two assumptions above. If own-account work is always available, even if the current task were to come to an end, in the subsequent period another one would be available. Because we take the return y, which fully characterizes the activity, to be an individual-specific parameter, the upcoming task is equivalent to a continuation of the previous one in every relevant aspect.

In intuitive terms, this assumption also translates the idea that one cannot be fired by himself. However, one can always decide to quit. In order to account for that possibility, we allow own-account workers to review their occupational decision at every period and pick the best option between looking for a job and working alone. Under those assumptions, we define the value of own-account work $OAW(y)$ as:

$$OAW(y) = \left(\frac{1}{1+\rho \cdot dt}\right) \cdot \left[y \cdot dt + \max\big(U, OAW(y)\big)\right] \tag{16}$$

At steady state, this expression is simplified further. When the parameters of the labor market are stable, if own-account work is preferred to job searching *at any point in time*, it will be preferred *at all points in time*. Thus, for any own-account worker, it must be that $\max\big(U, OAW(y)\big) = OAW(y)$ in all subsequent periods.[12] For this reason, we have that:

$$OAW(y) = \left(\frac{1}{1+\rho \cdot dt}\right) \cdot \left[y \cdot dt + OAW(y)\right] \tag{17}$$

$$OAW(y) + \rho \cdot OAW(y) \cdot dt = y \cdot dt + OAW(y) \tag{18}$$

$$\rho \cdot OAW(y) = y \tag{19}$$

[12] The reader may wonder why we bothered including the possibility of exiting own-account work if we are going to focus on the case where it never happens. We do it to stress that such permanence does not result from an a priori definition of own-account work as an absorbing state, but instead emerges naturally from the agent's sequential optimization.

5.1.5 The occupational choice

The usual job search framework assumes that, once the decision to enter the labor market is taken, individuals are either employed or unemployed. Here we allow workers to take into account what they can earn by themselves before looking for a job. Under this assumption, rational individuals will become own-account workers whenever the value of doing so is higher than the value of looking for a job:

$$OAW(y) \geq U \tag{20}$$

Equivalently, using the results from equation (15) and equation (19), this decision can be expressed as:

$$y \geq b + \frac{\lambda}{\rho + \delta} \cdot \int_{w_r}^{\infty} \left(w - w_r\right) \, dF(w) \tag{21}$$

The reader will notice that our discussion boils down to an expression analogous to the classic formulation for the participation decision, except that we assign a new interpretation to the outside option, that in our case is own-account work instead of inactivity. This is sufficient to motivate a set of implications for the prevalence of own-account work in the economy.

To see it, let us note that the share of autonomous workers in a given population is simply the proportion of individuals for whom the inequality above holds:

$$\mathbb{P}\left(y \geq b + \frac{\lambda}{\rho + \delta} \cdot \int_{w_r}^{\infty} \left(w - w_r\right) \, dF(w)\right) = \text{share of OAW in the workforce} \tag{22}$$

The intuition is that, given a sufficiently low reservation wage, even tasks with low productivity become attractive for a number of workers. To be precise, the comparative static analysis suggests that people are more likely to work on their own if:

1. *The return to own-account work is high enough.* Individuals with particularly high autonomous productivity are more likely to opt for own-account work. This result shows that the classic explanation, according to which people choose the occupation they are more skilled at, is indeed a particular case of our model.

2. *The unemployment benefits are low enough.* Lack of an insurance system decreases the value of the unemployment state.

3. *The arrival rate of offers is low enough.* Working alone is preferred when offers are too scarce anyway.

4. *Time discount rate is high enough.* When future earnings are discounted heavily, it becomes more important to secure an income source quickly.

5. *The destruction rate of wage jobs is high enough.* When jobs are short-lived, it is not rewarding to wait to get one.

6. *Expected wages are low enough.* Shifting the cumulative distribution of wages to the left decreases the expected return of looking for a job.

5.1.6 A time discount rate lower bound for own-account workers

Having established that equation (21) describes the occupational choice, we can take a step further and characterize it as a condition on the discount rate:

$$\rho \geq \frac{\lambda}{y - b} \cdot \int_{w_r}^{\infty} \left(w - w_r\right) \, dF(w) - \delta \tag{23}$$

To be clear, equation (23) is just an alternative expression for equation (21) that formalizes what is the minimum discount rate that ensures that the value of own-account work is higher than the value of unemployment, given the labor market parameters. This particular expression is of interest because it shows that, under realistic conditions, there can always be an arbitrarily high level of preference for the present that rationalizes the choice for own-account work.

We should note that the inequalities above are meaningful if we have non-negative and finite λ (the presence of frictions in the market); strictly positive $\rho + \delta$ (some preference for the present); finite expected wage and finite reservation wage (otherwise no finite outside option would beat it); and own-account work productivity strictly above the unemployment benefit ($y > b$, otherwise the inequality in equation (21) is trivially impossible).

5.2 What can we learn from the model?

Economic models are bound to be limited representations of the reality, but they can be instrumental if they can articulate important aspects of a given phenomenon, therefore providing guidance for the empirical analysis of it. Having presented the framework, now we can discuss how it addresses some relevant questions about own-account work at the theoretical level and, after that, how we operationalize it to estimate a lower bound for the share of financially constrained own-account workers in Brazil.

If anyone can work on their own, why is there unemployment?

The possibility of autonomous work does not make it necessarily a better choice than unemployment. The job search framework explicitly tells us that unemployment is a valuable state in itself because it creates the possibility of finding a job. Our extension builds on it to be specific about how the value of unemployment can be higher than that of an outside option such as autonomous work. Hence, it can be optimal for an individual to be unemployed if the returns on own-account work are low enough, if the unemployment insurance is high enough and if the discounted expected returns on wage employment are high enough, given the arrival rate of job opportunities, how long they last and how future income is discounted.

Why is own account work more prevalent in poor regions?

The comparative statics outlined in the previous section shows that the elements associated with a higher relative valuation of own-account work are also the elements commonly found in developing labor markets. To be precise, a high incidence of own-account work would be consistent with lack of social welfare policies (low b); scarcity of job vacancies (low λ); lack of long-lasting jobs (high δ), and underdeveloped financial markets and high interest rates (high ρ).

Why do people work on their own if there exist better-paid wage jobs?

The usual explanation suggests that individuals may value non-monetary aspects of own-account work. As an alternative to this view, our model suggests that a scarcity of job positions combined with a strong preference for present consumption is sufficient to justify this decision. For concreteness, the argument is that a family head with no savings may discount the future very heavily if day-to-day expenses are not secured. In that case, low earnings today are preferred to the alternative of investing time on the possibility of a minimum wage in 6 months, especially if the worker cannot access the financial market to smooth her intertemporal consumption. Importantly, if the labor market conditions are stationary, the individual will be stuck: at each moment she is allowed to choose differently, and at each moment she chooses the alternative that provides low, but immediate earnings.

If individuals decide by themselves to work alone, how can we talk about "necessity" self-employment? In which sense are they constrained?

The hypothetical case discussed just above illustrates how an occupation that provides a lower instantaneous income can still be more valuable for rational agents whose utility is determined only by monetary returns. In itself, it requires no market failure. If all transactions take place at market prices, there is little room to argue that their decision is not optimal in a strictly economic sense. The view from the dismal science is that workers facing low demand for their skills are as constrained as tourists who face high prices for ice-cream.

However, it can be the case that an individual who would like to anticipate future income *at the prevailing interest rates* in order to smooth her intertemporal consumption during the unemployment interval may be unable to do so because of failures in the financial market, be it missing markets or asymmetry of information. For this worker, the possibility to quickly receive some work income instead may become her best (constrained) option.

When we frame the problem this way, we can offer a more precise interpretation for the idea of constrained own-account workers as those for whom such an occupation is preferred to searching for a wage employment due to their high preference for the present, although they would have opted for looking for a better-paid wage job if they were able to finance consumption at the prevailing market rates during the search period. This approach has the benefit of backing the "necessity" occupational decision with a particular market failure, providing an objective meaning to the constraint.

5.3 Empirical estimation strategy

Our strategy to bring the model to the data is simply to translate the theoretical inequality on discount rates we established in equation (23) into its empirical counterpart. For that purpose, it is useful to reexpress the integral that appears in that expression as follows:

$$\rho \geq \frac{\lambda}{y-b} \cdot \left[\int_{w_r}^{\infty} \left(w - w_r \right) \cdot f(w)\ d(w) \right] - \delta \tag{24}$$

$$\rho \geq \frac{\lambda}{y-b} \cdot \left[\int_{w_r}^{\infty} w \cdot f(w)\ d(w) - \int_{w_r}^{\infty} w_r \cdot f(w)\ d(w) \right] - \delta \tag{25}$$

$$\rho \geq \frac{\lambda}{y-b} \cdot \left[\int_{w_r}^{\infty} w \cdot f(w)\ d(w) - w_r \cdot \left[1 - F(w_r) \right] \right] - \delta \tag{26}$$

$$\rho \geq \frac{\lambda}{y-b} \cdot \left[\mathbb{E}(w \mid w > w_r) - w_r \cdot \overline{F}(w_r) \right] - \delta \tag{27}$$

The important step here is to see how the difference between the stochastic job wage and the reservation wage integrated over the support of the acceptable wage offers is equivalent to the mean value of the acceptable wages minus the reservation wage multiplied by the probability the wage offer is above the reservation level.

After that, an empirical counterpart for equation (27) for a given own-account worker i, with a vector of attributes X_i, can be written as:

$$\hat{\rho}_i \geq \frac{\mathbb{E}\left(\lambda \mid X_i\right)}{y_i - \mathbb{E}\left(b \mid X_i\right)} \cdot \left[\mathbb{E}\left(w \mid w > w_r, X_i\right) - \mathbb{E}\left(w_r \mid X_i\right) \cdot \mathbb{P}(w \geq w_r)\right] - \mathbb{E}\left(\delta \mid X_i\right) \quad (28)$$

where the theoretical parameters are substituted by conditional expected values that can be estimated. We will discuss the details of each step in the next section, but the outline of the estimation plan is:

- $\mathbb{E}\left(w \mid w > w_r, X_i\right)$: The potential wage is estimated by fitting an OLS regression on the work income of paid employees;
- $\mathbb{E}\left(w_r \mid X_i\right)$: The reservation wage is assumed to be the bottom 10th percentile of the income observed for a worker with similar attributes X_i, and is fitted with a quantile regression;
- $\mathbb{E}\left(b \mid X_i\right)$: The unemployment benefit is taken to be the typical benefit reported by unemployed individuals;
- $\mathbb{E}\left(\delta \mid X_i\right)$: The expected job destruction rate is estimated using a parametric duration model of exponential form for the length of paid employment;
- $\mathbb{E}\left(\lambda \mid X_i\right)$: The expected job offer arrival rate is calculated using the estimated reservation wage and a parametric duration model of exponential form for the transition of unemployed individuals into paid employment;
- $\mathbb{P}(w \geq w_r)$: The probability of receiving an acceptable offer is estimated as the complement of the cumulative distribution function of a log-normal centered at the potential log wage and evaluated at the estimated log reservation wage;
- y_i: The income from own-account work is directly observed for individuals at this employment category;
- $\hat{\rho}_i$: Finally, we can calculate the discount rate threshold that is implied by the structural combination of the components above for all individuals that are observed at own-account work, for whom the inequality must hold.

The fundamental assumption we make in order to claim that equation (28) is a credible translation of equation (27) is that the estimation of those parameters is consistent with the agent's own perception of the market conditions they face. In other words, we take the econometric results to measure the empirical content equivalent to "how much people like me can make in a wage job?", "how many months is it going to take me to find one?" and "how long is this job likely to last?". At the core, we are just fitting an answer to each of those questions, in order to uncover how the people we observe as own-account workers are most likely to feel with respect to a particular unobservable factor, namely: "at which rate am I willing to change future consumption for present consumption?".

From a statistical perspective, we also assume that, conditional on the observable attributes, the disturbance terms of all the econometric models that enter into the structure above are exogenous and independent, which allows us to estimate them separately. Under this assumption, we do not require a correction for participation, for example, à la Heckman (1979). In section 8, we discuss alternative, less restrictive estimation strategies to be implemented in future work.

6 Empirical Results

This section presents the details of the estimation of each intermediary component of the structural model and examines their results. Our main objective here is to be transparent about how the individual attributes affect the conditional expectations that will be used to infer the lower bound of the individual discount rate.

6.1 Estimation of the expected wage offer

The first step of the empirical analysis is to estimate the income any given individual could expect to earn working in a wage job. In order to do so, we adopt a conventional wage regression in the tradition of Mincer (1974), assuming that the individual attributes we observe can lead to a conditional average that approximates the agents' perception of the job opportunities they could potentially access in the labor market. These results have a central role in the model because higher potential wages will make paid employment a more attractive option vis-à-vis own-account work, everything else constant.

We emphasize that neither the estimated potential wage nor the estimated reservation wage is required to be above the value of the official minimum wage defined for full-time jobs. There are two major reasons for this: the potential for employment to be a part-time job, what would imply a monthly wage that could be a fraction of the minimum wage; and the possibility of informal work, in which case the minimum wage is not enforced.

In the baseline estimation, the regression is performed on the log monthly income from the main job of all paid employees. This approach assumes that individuals build their expectations about potential offers based on the full distribution of wages, which is closer in spirit with the theoretical model since we do not allow for returns to seniority or on-the-job search. However, as a robustness check, we repeat the estimation using only the wages of those individuals who are known to be employed in a wage job for less than 12 months. The idea is that the work income reported by the recent hires can be closer to the available vacancies and, as such, could be the relevant guidance for those considering whether or not to look for a job. The coefficients for both models are reported in table 3 in the appendix.

In choosing the covariates, our objective was to be flexible, yet parsimonious. In all regressions, we have split the age into discrete levels and interacted those levels with gender in order to allow for non-linearities in those dimensions. In the baseline specification, we observe that the wage is generally increasing in age, but more so for males. A female worker with age between 40-44 can expect to earn 72% more than at the very beginning of her career. In comparison, a male worker also between 40-44 has on average a wage 158% higher than the same comparison group.

The classic human capital measurement, schooling, is also included in the estimation in discrete levels, which are interacted with race, in order to capture differences in returns to education due to labor specialization preferences or potential discriminatory issues. Wage is increasing in schooling, but less so for non-whites (i.e. self-reported blacks, browns, and indians). A non-white individual with high school can expect to receive a wage 47% higher than a non-white worker with no formal education, while whites with high school can typically earn 60% more vis-à-vis the same reference group. As expected, there is a strong return to the ability to read and write, commanding a 45% premium over illiterate workers. People still at school generally earn less, which is consistent with the practice of internships or part-time work by this population.

Moving to the dynamics of local markets, we find that urban areas offer, on average, 16% higher wages than rural areas. The last component of the estimation is the regional dummies, which are assigned to each one of the 27 federal divisions of Brazil, with separate indicators for the capital and for the rest of the State. In line with the evidence presented in figure 3, the industrialized states in the southern half of the country, such as São Paulo (SP), Rio de Janeiro (RJ), Paraná (PR) and Santa Catarina (SC), all offer significantly higher average wages, particularly in their capitals, where the positive gap versus the hinterland of Rondonia (the omitted category) is at least 25%.

6.2 Estimation of reservation wages

Given that the PNAD survey does not pose a direct question about the lowest wage the individual would be willing to accept, we need to estimate it. In principle, one could simply take the absolute lowest value observed at the conditional cells defined by a set of individuals attributes. The main drawback of this option is that it is extremely vulnerable to outliers, and would also require sufficiently large cells to be consistent, as the estimation for a minimum converges much more slowly than the estimation for the average.

To overcome those issues, our strategy is to use quantile regression to predict the conditional expected value *at a sufficiently low rank in the wage distribution.* In our baseline specification, we assume that the 10th percentile of the distribution is a reasonable proxy the reservation wage. To examine the sensitivity of the results to different cutoffs, in the robustness analysis we replicate the estimation with 5th and 15th percentiles.

The most important difference with respect to the previous estimation is that now we introduce family characteristics. This econometric choice, which serves as an exclusion restriction, is motivated by the prior that living alone or having two kids should not affect the wage opportunities a worker expect to see in the market directly, but it could affect the minimum monthly income she is willing to accept.

This intuition is borne out by the data, as reported in table 5. The 10th percentile wage of workers who are the head of the household and have no partner is found to be about 6% lower in comparison to heads with a partner. Furthermore, this value decreases by about 2% to 3% for each additional household member under 24 or above 65 years of age. We interpret these results as evidence that, everything else constant, being a single parent or having many dependents is associated with a lower reservation wage. This result is consistent with a higher preference for a part-time job (hence, with lower monthly payments), which limits the time outside the house, but also with a lower selectivity regarding offers, due to pressing family consumption needs.

The signal of the remaining coefficients is largely aligned with what we found in the previous section, although the marginal effects there affect the average wage, while here they change the expected wage at the 10th percentile rank. As expected, the bottom of the distribution of wages for more educated individuals is significantly above the corresponding values for uneducated workers. One interesting distinction is that the reservation wage for females peaks between 25 and 34 years old, versus 40-44 for man, and then decreases for older workers. Moreover, the gains from living in a city are much stronger for the bottom of the distribution (+39%) than for the average (+17%), which is likely a consequence of stronger enforcement of the minimum wage.

The implication for the occupational decision is that the reservation wage is the key benchmark against which the own-account productivity is compared: if the minimum acceptable payment is relatively low, then it is easier for the worker to exceed it working alone.

6.3 Estimation of transition hazards in and out of paid employment

So far we have estimated the wage someone can expect to receive as a paid employee and the typical values that are found at the lower end of the distribution of accepted wages. However, paid jobs are not found instantaneously and do not last forever. In order to calculate the value of looking for a job opportunity, we also need to estimate how long people usually spend in unemployment and how long those jobs typically last.

Those questions have a direct empirical measurement with duration analysis.[13] Because our theoretical model assumes agents forming expectations at steady state, a consistent choice is to use a parametric duration model that fits the duration outcome using an exponential distribution and, by construction, estimates transition hazards that are time-independent.[14]

In order to operationalize the empirical content of entry into paid employment as closely as possible, all other transitions from unemployment (namely, into inactivity or into self-employment) are treated as censoring — intuitively, those changes prevent us from observing a transition into a wage job, in the same way that the end of follow up does. In other words, we are interested in a risk-specific hazard. On the other hand, in the case of end of employment, we treat all transitions as an event of job destruction, since the present discounted valuation

[13] Here we follow a long tradition in applied labor economics that uses duration (or survival) estimation techniques to model individual spells in different employment states, in the spirit of the classic works of Kiefer (1988) and Meyer (1990). For a comprehensive treatment of these techniques, see Kalbfleisch and Prentice (2002). For other applications of duration analysis in the Brazilian context, see Menezes-Filho and Picchetti (2000) and Margolis (2009).

[14] An alternative parametrization using the Weibull distribution, which is a generalization of the exponential case that allows for monotonically increasing (or decreasing) hazards, suggests that the job exit rate actually decreases over the length of the spell, meaning that people usually quit jobs at a higher rate at the initial months of employment. Still, the estimated Weibull parameter is 0.84, reassuring that the constant hazard is not a far fetched simplification. In the case of the duration of unemployment, the approximation is even more convincing, as we cannot reject a unitary Weibull parameter and, consequently, a constant hazard of transition into employment. Those results are available in table 7 at the appendix.

of the job is affected only by its expected duration, not by the subsequent state.[15] In both cases, we assume transitions to take place at the midpoint of the interval between the last time the individual is observed at the original state and the first time she is observed at a different state (typically a window of three months, which is the interval between the survey rounds).

As reported in table 6, there is little variation on the transition rates for different household roles, with the major exception of those identified as children, who have higher unemployment duration and lower employment duration. As in the previous regressions, there are clear distinctions in age and gender: it is possible to identify a prime period for employment, around the 30s for men and 40s for women, when the expected duration of unemployment is the lowest and the job are expected to be the most stable. Across all ages, women generally spend more time looking for a job, but also stay employed for longer.

Putting the two estimations side by side, it is interesting to see that living in a city and having higher levels of education are associated with longer spells at *both* states. In other words, uneducated individuals and those in rural areas are observed to jump between jobs of short duration, while college graduates and those in urban centers typically wait longer, but also get more stable jobs. This pattern complements and reinforces the previous results describing higher average wages and higher reservation wages for urban workers with higher human capital.

Finally, looking at the regional heterogeneity, we find that the distinction is more pronounced with respect to the duration of employment, with rich areas having a lower job destruction rate than the poor regions.

An important regularity that emerges in this estimation is that some characteristics that are associated with a shorter paid employment spell (men without college, living in rural areal at the northern region of the country) are also disproportionately present among own-account workers, which is in line with the models' predictions.

6.4 Estimation of the job offer arrival rate

It is important to note that the λ parameter from the theoretical model represents the *rate of arrival of job offers*, which is not the same as the *rate of transition into a job* estimated in the previous section. The first represents the frequency according to which new employment positions appear to the job-seeker, while the latter accounts only for the opportunities that are sufficiently attractive to be accepted.

[15] This could change in a model of "stepping stones", where being at a job increases the chances of moving into another job. We abstract from this dynamic here.

By its own nature, the arrival of job offers is much harder to track and is usually not included in labor surveys. However, we can use the reservation wage defined above to infer the job offer rate from the estimated job transition rate, under some distributional assumptions.

To be clear, let h denote the hazard of the transition into employment (i.e. the one we estimated with the duration model). As before, λ and w_r represent the job offer arrival rate and the reservation wage, respectively. The relationship between those elements is given by:

$$h = \lambda \cdot \mathbb{P}(w > w_r) \tag{29}$$

Hence, to calculate λ, we need to recover $\mathbb{P}(w > w_r)$. Assuming that the distribution of the potential wage offers for a given individual is a log-normal centered in the expected log wage we estimated in section 6.1 (say, $\hat{w}$) and with a variance analogous to the empirical variance of the fitted wages ($\hat{\sigma}^2 = \mathbb{V}(\hat{w})$), then we have that:

$$\mathbb{P}(\hat{w} > \hat{w}_r) = 1 - \Phi\left(\frac{\hat{w}_r - \hat{w}}{\hat{\sigma}}\right) \tag{30}$$

Under the baseline specification, we estimate that the conditional unemployment duration for the Brazilian own-account workers would be about 5.4 months (given a 0.185 transition rate). Using the strategy above, we find that 95% of the offers would be acceptable, on average, which implies one offer every 5.2 months (or a 0.189 arrival rate).

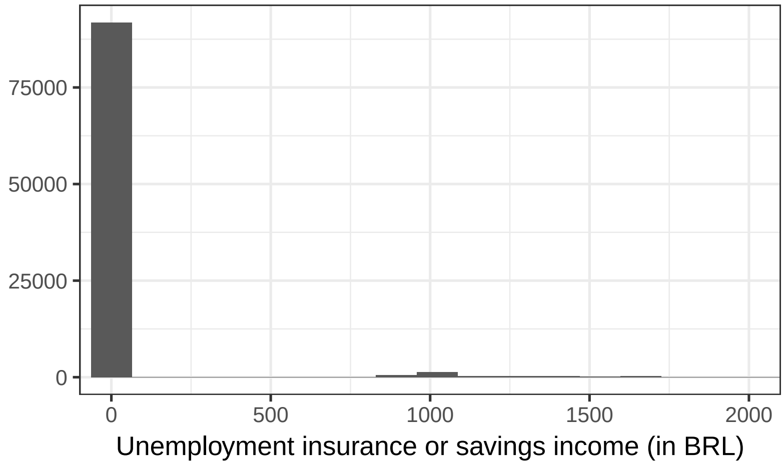
6.5 The expected value of unemployment benefits

The theory anticipates that a strong insurance system would increase the value of looking for a paid job. In the case of Brazil, the information available at the PNAD survey suggests that such form of income is negligible — most often, job-seekers report receiving no insurance at all, while a very small share has benefits around the minimum wage, as shown in figure 6.

Two factors contribute to the weak role of unemployment insurance in the Brazilian context. First, the benefit can be provided only in the case of unjustified layoff of formal employees. Second, it lasts for at most five months, which is a bit less than the average unemployment duration.

Given the very large and unsystematic presence of non-insured job-seekers, there is very little informational gain from estimating the probability of receiving the benefit (using probit or logit techniques, for instance) or from modeling the expected value of the insurance using the usual covariates in a standard OLS.

Figure 6: *Histogram of unemployment benefits and savings income for unemployed individuals. Source: Own calculation using PNAD microdata.*



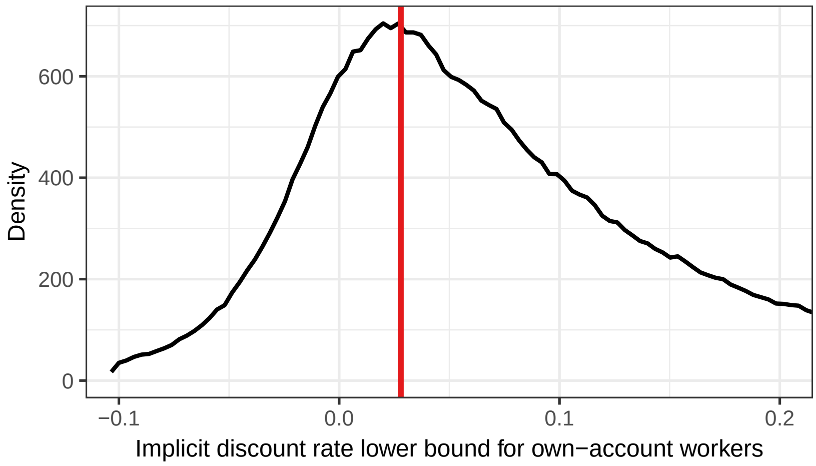
For this reason, we simply assume that workers do not account for any insurance when choosing their occupation. This strategy has the appealing feature of being trivially true for all individuals looking for a job for the first time, and it is also realistic for the largest share of job-seekers in any unemployment round.[16]

6.6 Main results: the subjective discount rate lower bound

Now we have all the pieces to calculate the implicit time discount rate of the Brazilian own-account workers, as defined equation (28). Rigorously, the object we recover is a *lower bound for* ρ, in the sense that it is the minimum discount rate that makes the value of own-account work higher than the value of looking for a wage job, given the potential labor market conditions the workers are facing. Since this inequality must hold for every rational agent who has revealed their choice for working alone, we can calculate the individual ρ lower bound and examine the distribution of such minima in the population of own-account workers, whose empirical density is plotted in figure 7.

[16] We have considered extending the analysis to an extreme scenario where all agents take for granted they would be entitled to the standard benefit during unemployment: three payments at 80% of their previous wage, with a floor at the minimum wage and a cap that is updated every year by special legislation. However, this specification leads to a significant volume of cases where the benefit is higher than the observed own-account income, a situation that is theoretically and empirically implausible. Potential validation strategies for our approach are discussed further in section 8.

Figure 7: *Empirical density of the implicit discount rate lower bound for Brazilian own-account workers at baseline specification. The reference line marks the average credit rate for individuals in 2016-2017 (2.8% per month). Sources: Own calculation using PNAD microdata and Brazilian Central Bank.*

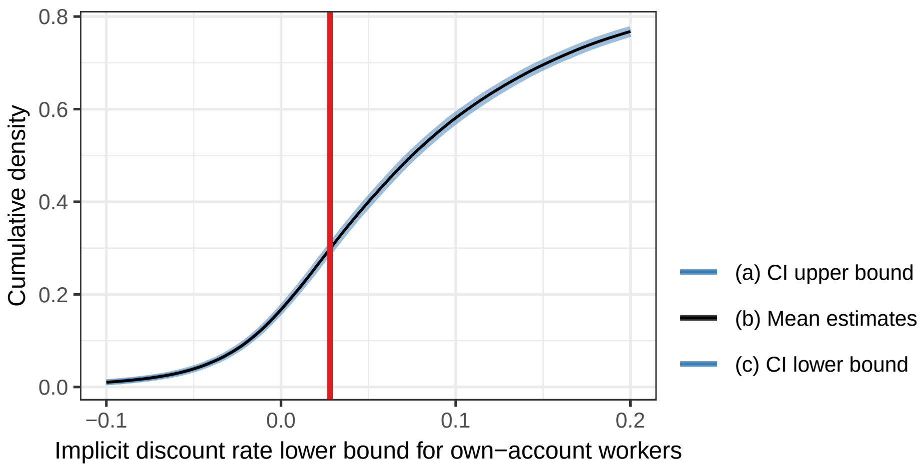


The profile of the minimum subjective discount rates estimated for the population of own-account workers has some interesting properties. First, it is asymmetric, with a heavy right tail, indicating that the incidence of very high minima for ρ is more frequent than the incidence of very low minima. Second, it is not bound by zero, but it does not mean that people are actually behaving according to a negative rate — since it is a lower bound, we cannot be sure of their actual discount rate, we can only say that individual rationality requires it to be somewhere above that estimated level. Finally, the mode is close to 2.5%.

But what does it mean to discount the future at 2.5% per month in Brazil? To make sense of this finding, we compare it to the interest rate priced in the credit transactions actually observed between banks and individual borrowers. For simplicity, we focus on the average credit rate weighted by the volume of new operations (mostly mortgages, payroll lending, credit card, vehicle financing, and consumer credit), as reported by the Central Bank. Over the period of analysis (2016-2017), the average for such interest rate was 2.8% per month.

This reference is useful because it suggests that our strategy did recover results around realistic levels. More importantly, it can serve as a threshold: when the lowest rate consistent with the observed occupational decision of the own-account worker is strictly above the average market rate for personal credit, we have some indirect evidence that the agent did not have access to such market, because had she been able to borrow at the prevailing rate, she would have chosen to search for wage work. To measure how common is this case, we can study the cumulative density of the lower bound of ρ, as plotted in figure 8.

Figure 8: *Empirical cumulative density of the implicit discount rate lower bound for Brazilian own-account workers at baseline specification. The confidence intervals at 1% level were obtained with clustered bootstrap over 400 replications. The reference line marks the average credit rate for individuals in 2016-2017 (2.8% per month). Sources: Own calculation using PNAD microdata and Brazilian Central Bank.*



At the lower end of the graph, we have that 20% of the Brazilian own-account workers have a return that is sufficiently above their potential paid job opportunities to ensure that their occupation would remain preferable even in a hypothetical context of zero interest rate. In other words, autonomous work offers a better value for them regardless of any possible preference they might have for the present.

There is an intermediary group, with about 10% of the own-account workers, whose minimum implicit discount rate lies between zero and the market rate. In this case, their option for own-account work is driven by some degree of preference for the present but there is no evidence of a constrained choice.

Finally, we find that nearly 70% of the own-account workers in Brazil appear to be in this occupation due to a financial constraint, in the precise sense that the lowest time discount rate implied by their choice is strictly higher than the prevailing interest rates.[17] In this case, the instantaneous return they have as own-account workers are so low (relative to potential labor market opportunities available for individuals similar to them) that rationalizing their decision requires a particularly strong time discount rate — in fact, so strong that they would have preferred to finance present consumption and search for a job, if they had access to the market rate.

[17] It is interesting to note that each piece of the model we estimate has, in general, very precise coefficients. That explains why the bootstrap confidence intervals are so narrow over the whole CDF: any imprecision due to sampling is marginal due to the size of the sample available.

If our results are representative, we can argue that traditional models of comparative advantages, without frictions and without time, would be a fair representation for only 20% of the own-account employment in Brazil. This is important because, in the majority of the cases, the intertemporal trade-off appears to be sufficiently relevant to drive the occupational choice. On top of that, their high preference for the present suggests that those workers are facing liquidity stress and are not able to finance consumption at the market rates.

This evidence is consistent with the hypothesis that a share of the own-account workers in developing markets are making a labor supply decision that is not their first best under available market prices. Instead, their choice reflects high frictions in the labor market and imperfections in the financial market. It is in this context that the possibility to bypass the job search process makes own-account work the best *constrained* option. At the core, our main finding is an evidence of market failure, suggesting the possibility of improvement in the allocation of resources.

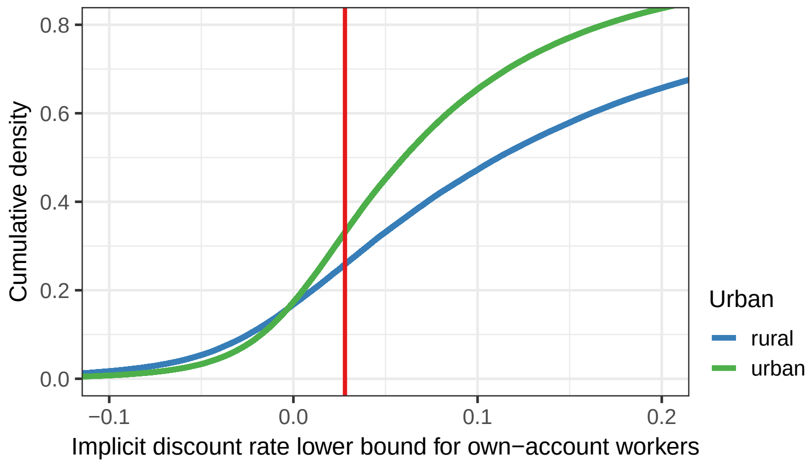
This interpretation is reinforced by the conditional distribution of the lower bound of ρ by groups, presented in figure 9. For instance, 74% of the own-account workers in rural areas appear to be financially constrained, compared to 67% in urban centers, a result that is coherent with the presence of more developed financial agents in cities. This finding signals that urgent consumption needs combined with imperfect credit markets do play a role in the high prevalence unskilled own-account work we observe in the hinterlands of the country.

Next, we examine the distribution of the implicit discount rate lower bound according to gender. In all estimations that enter the model, we have found statistically significant differences — males have better wage offers and lower unemployment duration, females stay in the job a bit longer — and yet they have a similar distribution of minima discount rates. Indeed, if one admits that males and females have no intrinsic reason to differ in their intertemporal preferences, this result reassures the internal validity of the model.

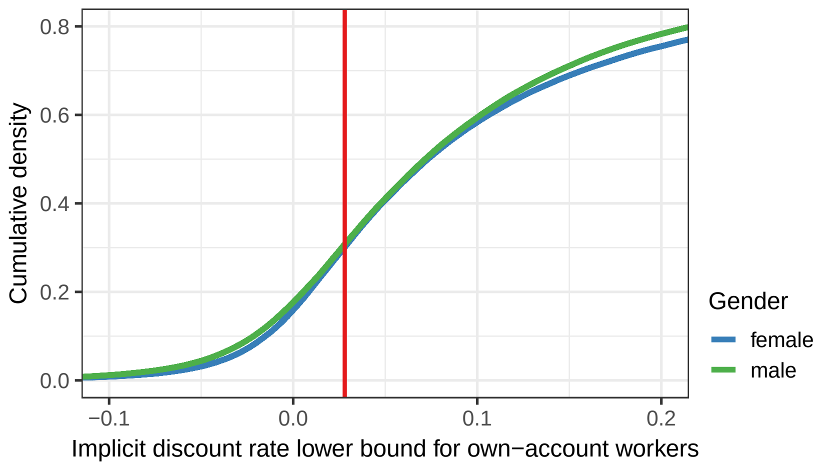
Moreover, beyond the gender, the level of responsibility within the family appears to be related to the individual consumption urgency. As intuitively expected, family heads have the highest time discount and are the ones most often constrained (73%), followed by partners (69%), and then children (56%). These results suggest that financial constraints push family heads into own-account work, while partners are able to choose somewhat better own-account work opportunities when they decide to do so.

Figure 9: *Conditional distribution of the implicit discount rate lower bound for Brazilian own-account workers at baseline specification. The reference line marks the average credit rate for individuals in 2016-2017 (2.8% per month). Sources: Own calculation using PNAD microdata and Brazilian Central Bank.*

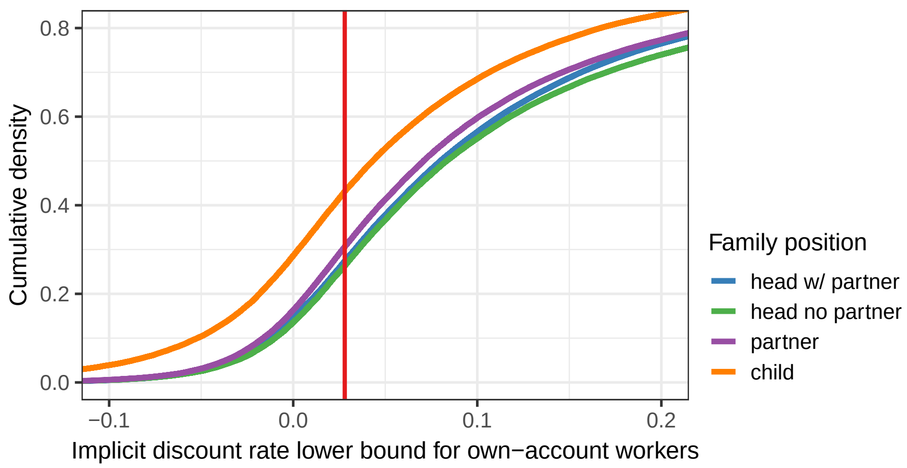
(a) *Split by urban status*



(b) *Split by gender*



(c) *Split by family position*



7 Robustness Checks

Our baseline specification relies on a set of particular assumptions that are necessary to operationalize the theoretical model. In order to increase the transparency of our estimation, now we discuss how different choices would affect the results.

7.1 All workers versus new hires as the reference for the potential wage

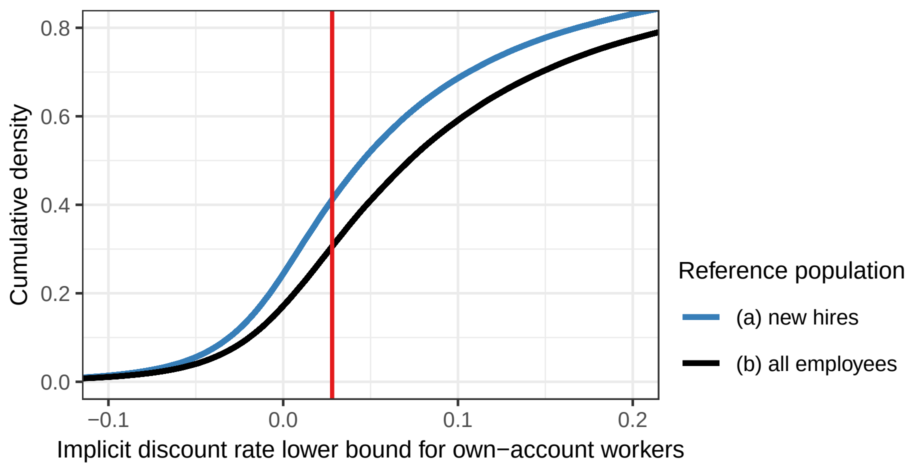
In order to calculate the value of the potential job opportunities faced by own-account workers, we have predicted their expected wages (and their expected reservation wages) using the coefficients obtained from a linear regression (and a quantile regression) on the work income reported by the full population of employees. This strategy should be preferred if people build expectations simply by looking at how much similar workers are making.

Alternatively, we can estimate the counterfactual wages using only the remuneration reported by new hires — employees who are in the job for less than a year. This strategy should be preferred if the relevant reference in the decision between own-account and paid employment is the potential entry wage.

When we do so, the median potential wage decrease by 23% and the median reservation wage decreases by 38% since the population used as a reference now excludes people with longer job spells, who usually have higher earnings. Looking at table 3, table 5 and table 4 we can see that this difference is driven primarily by lower coefficients on the returns to age and to education.

Because this alternative specification leads to lower potential wages, a lower time discount rate is required to make own-account preferable, everything else constant. As a consequence, the estimated share of constrained own-account workers falls by 10 percentage points, from 69% to 59%, as shown in figure 10.

Figure 10: *CDF of the implicit discount rate lower bound for Brazilian own-account workers, using as reference the wage of all workers (baseline) and of new hires. The reference line marks the average credit rate for individuals in 2016-2017 (2.8% per month).*



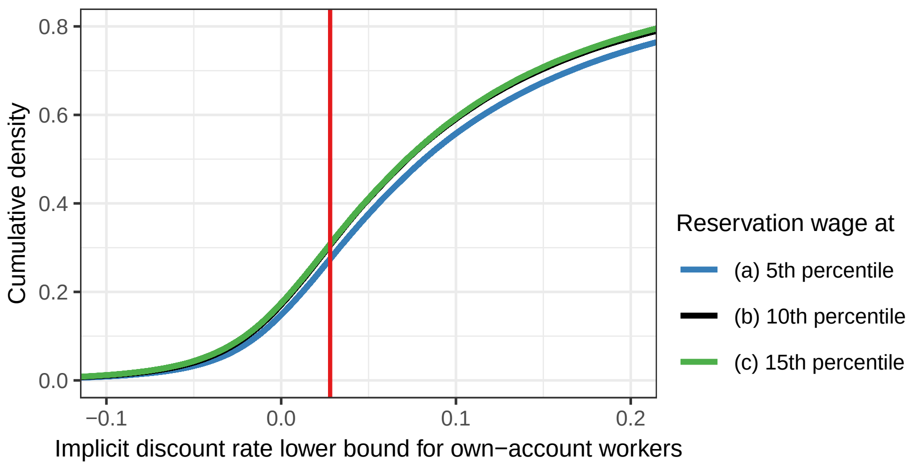
This is the most important alternative specification of our study — the difference is meaningful and we have no strong reason to rule out the plausibility of this result. In any case, it is instructive that, even under a more conservative measurement of potential wages, we still find that 6 in 10 own-account workers appear to be financially constrained.

7.2 Alternative quantiles in the estimation of the reservation wage

Following a similar rationale, we can examine different choices for the percentile cutoff that we used to approximate the reservation wage. As discussed above, we aim to estimate the wage level below which the job offers are most likely to be rejected, conditional on the observable attributes of the worker, but the baseline reference at the 10th percentile is arguably arbitrary.

In this respect, it is reassuring to find only a modest variation in the results under different cutoffs, as plotted in figure 11. In particular, the use of 5th percentile (the blue line) leads to 4 percentage points increase in constrained own-account workers, while the results with a 15th percentile (the green line) are nearly indistinguishable from our baseline case. The main reason behind this asymmetry is that the distance between the fitted reservation wages in the 5th percentile specification and the baseline case is larger than the gap between the baseline and the 15th percentile specification.

Figure 11: *CDF of the implicit discount rate lower bound for Brazilian own-account workers, estimating the reservation wage as the 5th, 10th (baseline) and 15th percentiles. The reference line marks the average credit rate for individuals in 2016-2017 (2.8% per month).*



7.3 Matrix of alternative specifications

The exercises above have shifted one assumption each time, now we present the potential results from all combinations of the alternative specifications. The most extreme case is obtained with a very low cutoff for the reservation wage and the use of the full sample of employees, in which case 73% of own-account workers would be constrained, while the polar case suggests 58%. In summary, our results offer some confidence that, under a set of specifications, the estimated share of constrained own-account workers lies within the 60-70% interval.

Table 2: *Estimated share of constrained own-account workers under different specifications*

	Reservation wage at 5th percentile	Reservation wage at 10th percentile	Reservation wage at 15th percentile
Reference sample: All employees	0.73	0.69	0.69
Reference sample: New hires	0.64	0.59	0.58

Notes: The robustness checks covers two dimentions (1) a variation on the assumption over the distribution of the wage offer, using data from (a) all workers or (b) only from workers with less than 12 months in the job; and (2) the different cutoffs assumed in the reservation wage estimation, at the 5th, 10th and 15th percentiles.

8 Conclusion, Caveats and Extensions

This study considered how the interaction between labor market frictions and financial markets failures can help to explain the higher presence of own-account work in developing countries. Absent frictions, workers can simply pick whatever occupation offers the highest income, but if job-seekers take time to find vacancies and cannot finance their present consumption, there is a trade-off between a securing a small return quickly and waiting for higher income in the future. Using a simple extension of the job search model, we showed how a sufficiently high discount rate can make own-account work more valuable than unemployment.

In the empirical application, we used the counterpart of the model to infer how high such discount rate needs to be for the particular case of own-account workers in Brazil. We find that 69% of them appear to be financially constrained, as their occupational choice suggests a discount rate that is strictly above the market rate. Additionally, our results highlight the importance of allowing for heterogeneity in time discounting, a parameter that is most often assumed homogeneous (and usually at the same level of the market rate) in theoretical and empirical research.

8.1 Consistency of the identification

The identification strategy adopted in this research assumes that the piece-wise estimation of the structural parameters of the model is both consistent and sufficiently close to the agents' perceptions of the labor market conditions that inform their occupational decision. Importantly, our findings would be biased if (a) the structure we assume fails to capture the relevant determinants in the heuristic decision-making process or (b) the results we find as econometricians do not reflect the beliefs and perceptions of the agents.

In the first case, we might have imposed an incorrect structure or omitted crucial elements in the agent's labor supply decision. This would be the case if social norms and peer effects are the major determinants, making workers less responsive to pure monetary returns. One could imagine, as an alternative hypothesis, that workers in poor regions are particularly averse to hierarchy and therefore more inclined to work on their own, regardless of their time discount. As it often happens with preferences, such hypothesis could only be supported with a strong justification for why this particular taste would be more prevalent in poor countries, which we do not have at this point.

In the second case, should the agents be systematically more (or less) confident on their chances of finding a job, or about the level of wages that firms are offering, compared to what our estimates suggest, we would overestimate (or underestimate) the implicit discount rate threshold that is consistent with their choice of own-account work. We cannot rule out that the agents are systematically mistaken about their own opportunities, but this possibility would require some form of limitation on their rationality. Alternatively, *we* might have mismeasured the worker's potential opportunities, particularly due to omitted variables, which we discuss below.

8.2 Estimation refinements

From a statistical perspective, we assumed conditionally independent errors between the pieces of the model. This restrictive assumption could be relaxed in future development of this work if we estimate simultaneously all four equations and allow for a correlation structure among the unobservables.

A more straightforward refinement would be to estimate the two duration models acknowledging the interval-censored nature of the data. Here, we adopted the usual simplification that assumes a transition in the midpoint of the interval between the date the individual is observed in the initial state and the following date when he is observed in a different state. Because we have quarterly observations, this simplification may be somewhat restrictive and could be avoided with a likelihood maximization that accounts for interval observation.

Regarding the estimation of the reservation wage, Hausman et al. (2019) show that the quantile regression estimator could be biased if the observed wages are measured with error. Future work could apply their sieve maximum likelihood estimator the get a better estimate of the conditional quantile in the presence of measurement error. An alternative approach to be examined in future research is to focus exclusively on full-time workers or to account for differences in worked hours explicitly and transform all measured income into a full-time equivalent, using the 44 hours per week established in the Brazilian labor market. That would eliminate the differences due to workload, which would make it more plausible to explore estimators for extremal quantiles (Chernozhukov, 2005).

8.3 Quality of data

The use of a more transparent reservation wage, ideally reported by the individual herself, would improve the credibility of the estimation. To the best of our knowledge, there is no survey that captures this information systematically in Brazil.[18]

Second, it would be interesting to investigate further the issue of unemployment insurance, potentially leading to a strategy to distinguish with more confidence those individuals that are most likely to be eligible. This could be achieved with administrative data, that is more precise than the PNAD survey for the kind of information required to calculate the eligibility and the value of the benefit. In any case, the examination of a different data source could provide more confidence in the validity of our zero-insurance assumption.

Another potential limitation is that we lack detail on human capital. The data currently available for Brazil does not distinguish between an individual with a Ph.D. in Medicine and a Bachelor in Business Administration. It is not clear the extent to which the lack of area of specialization could have biased the result, but any imprecision introduced by it is likely to be small since the majority of the workers do not have a college degree at all.

[18] In lack of better data, an alternative estimation strategy would be to assume that the wage offers follow a known distribution with a tractable closed form, which would allow us to isolate the reservation wage in equation (15) as a function of the remaining parameters. The challenge here lies in finding a distribution that is simple enough and that offers a credible approximation to the empirical distribution. After a set of exercises, the best alternative we found was the Pareto distribution, but we believe its sharp discontinuity at the left side is too strong to accommodate the data.

More importantly, there is very limited visibility on the past experience of the worker, which prevents us from controlling for the potential accumulation of category-specific skills, a piece of information that is taken into account by the individual when assessing her options. We anticipate that adding appropriate control for past experience would narrow somewhat the gap we found between the estimation using new hires and all employees.

8.4 Alternative interpretations

The most immediate objection to our approach is that own-account workers might be willing to receive lower pay because they enjoy what they do. If that is systematically true, our result overestimates their discount rate and, consequently, the share of constrained cases. While we do not discard the possibility of preferences for occupations, we are skeptical that they could be the main driver of our results because the bulk of the own-account workers in the Brazilian context are at such low levels of income that any gains in work satisfaction are likely to be of second order with respect to earnings considerations.[19]

Second, we note that part of the effect can be driven by debt aversion, which is absent in our model. If workers have some degree of distaste for credit, they can take economic decisions based on a subjective interest rate persistently above the market rate, even though such a rate *is indeed available to them*. In this case, the effect for the occupational decision would be the same: debt averse individuals with high subjective discount rates would prefer own-account work, the same way credit constrained ones do. However, the fundamental cause would be different, since debt aversion implies *suboptimal demand* for credit while credit constraint is a question of *suboptimal supply*. This distinction would be relevant to public policy because it requires different policy tools. For instance, the implementation of subsidized credit lines for the unemployed would be much more effective in the presence of a pure financial constraint than in a case of debt aversion. Furthermore, debt aversion could explain why reductions in the market rates may not necessarily be followed by a drop in own-account work even when credit is available.

[19] In the developing context, the opposite might be true: own-account workers may *dislike* their occupation. Banerjee and Duflo (2011) describe that the most common aspiration of poor parents around the world, many of them working on their own, is for their children to grow up to get a salaried job, preferably a stable position in a government office. Such perspectives suggest caution when assuming widespread entrepreneurship aspirations in the Brazilian context.

8.5 Policy implications

Our finding adds a layer of complexity to the usual consensus in development economics according to which credit constraints *prevent people from working on their own*, since we claim that credit constraints actually *prevent people from finding a better-paid job*. This apparent contradiction is due to the confusion over the empirical content of self-employment, as we discussed in section 2. If one has in mind that entrepreneurial activity requires capital (as in the case of a potential employer), then credit constraints hinder self-employment. In our model, by contrast, own-account work is valuable because it offers the possibility to bypass the job search period, thus the usual credit constraints will foster it.

That said, the immediate policy implication of our model is that *consumer credit* availability could increase participation in the wage employment market. We are proposing policies that make it possible for job-seekers without savings to smooth the drop in consumption during their search period. This is distinct to the traditional earmarked microcredit policy, which is most often tied to an investment in the business itself. In our framework, directed microcredit and training foster own-account work through an effect on y, the autonomous productivity; while the availability of consumer credit would foster wage employment through a potential decrease in ρ, the relevant discount rate of the decision process.

We are able to claim that this would be welfare improving because the majority of own-account work in Brazil appears to be a misallocation of labor force driven by a combination of market failures. However, the most efficient way to address those market failures remains an open research question. It would be important to investigate whether a policy of unearmarked credit for the unemployed would be superior to the adoption of a comprehensive unemployment insurance rule, including those looking a job for the first time, from the point of view of public finance, banking stability, credit risk, and social welfare. A serious examination of it, however, requires equilibrium models, as we discuss below.

8.6 Extensions

The first important simplification of the model is the assumption of stationary parameters. Among other things, this approach disregards on-the-job skill acquisition and career promotions. Because earnings appear to increase faster (and for longer) in wage jobs, we might be underestimating its discounted value when we characterize the job as a fixed wage. At the same time, two factors suggest this could be a fair approximation. If individuals are myopic, they will decide based on the offered wage, not on the full, uncertain career path it might lead to. Second, given the very high discount rates we identified for those who are choosing own-account work, the discounted value of the possibility of a career progression 10 years in the future will add very little to the present value of a job.

The model also assumes away any issues related to physical or financial capital. As we described in section 4, the profile of the typical own-account worker (limited education, lives in rural areas and often receives less than the minimum wage) suggests that such occupation does not require a large accumulation of capital. The decision to become an employer, by contrast, is likely more dependent on capital considerations. An interesting extension of this model would be to integrate the employers, whose return can be seen as a function of their productivity, as we did with own-account workers, combined with the productivity of their invested capital and of their hired workers. Then, as in Evans and Jovanovic (1989), the lack of collateral might prevent skilled employers to fund the optimal amount of capital required by their productivity. Ultimately, future research could formalize in a common framework how the presence of liquidity constraints can, at the same time, hinder employers and foster own-account workers.

A related limitation is that we take an approach of partial equilibrium. The main drawback of having exogenous wage offers is that we cannot propose counterfactual exercises with the parameters we estimated. For instance, a general equilibrium model could allow us to infer by how much the share of own-account work in Brazil would fall if the government were to enforce universal unemployment insurance, or if online job search were to increase the offer arrival rate. Our limitation is that, in equilibrium, those counterfactual shifts would affect the distribution of job offers, an effect we cannot capture. A general equilibrium approach could also incorporate the consequence of competition within own-account workers, and between them and the firms, addressing questions such as: how elastic is the demand for own-account work services? Are they substitute for the goods and services offered by firms, or do they constitute a separate market?

Furthermore, in the same way we omitted preferences for a particular occupation, we also omitted any disutility of labor: in our model, the monthly income is sufficient to characterize the employment position. In practice, it means that a job that pays 1,000 for 40 hours is identical to a second one that pays 1,000 for 20 hours, and is preferable to a third one that pays 500 for 20 hours. This simplification is appropriate if individuals are mainly targeting a monthly income in their labor supply decision, which is the instance we take in this study and the one that makes the discussion about time discount rate the more straightforward. However, a more complex approach would be to take an employment opportunity as a bidimensional object, defined by a combination of the number of hours demanded *and* the hourly pay.

Such an extension, incorporating labor disutility and work hours, is particularly important to examine the trade-off between consumption vs. family-time. Workers with many dependents may prefer own-account work due to its quick returns (as we stressed in this study) but also because it opens the possibility to fine-tune the desired labor supply, in contrast to the usually limited flexibility of the full-time and part-time work contracts. By the same rationale, this extension would allow us to discuss the large presence of older individuals among the own-account workers, from whom this occupation can serve as a "bridge" into retirement, one that allows for a gradual reduction in work hours, as suggested by Ramnath, Shoven, and Slavov (2017) and Ameriks et al. (2017).

To conclude, we note that, at the empirical side, it would be important to validate our set of implications independently with reduced form estimations. In particular, the study of an exogenous variation in credit access for a well-defined community could provide further evidence of the direction and magnitude of its effect on the prevalence of constrained own-account work. Alternatively, future research can investigate the determinants of the subjective time discount, which can lead to better identification of its effects on occupational decisions. For instance, we suggested that single parents could have a stronger urgency for present consumption — a new policy for daycare would likely alleviate some of it, without affecting the discount rate of individuals without kids. If that is the case, a decrease in own-account work for that particular population would have a causal interpretation that is coherent with our results.

References

Albrecht, James, Lucas Navarro, and Susan Vroman. 2009. "The Effects of Labour Market Policies in an Economy with an Informal Sector". *The Economic Journal* 119 (539): 1105–1129. doi:10.1111/j.1468-0297.2009.02268.x.

Ameriks, John, et al. 2017. *Older Americans Would Work Longer If Jobs Were Flexible.* Working Paper 24008. National Bureau of Economic Research.

Banerjee, Abhijit V, and Esther Duflo. 2011. *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty.* New York: PublicAffairs.

Baptista, Rui, Murat Karaöz, and Joana Mendonça. 2014. "The Impact of Human Capital on the Early Success of Necessity versus Opportunity-Based Entrepreneurs". *Small Business Economics* 42 (4): 831–847. doi:10.1007/s11187-013-9502-z.

Blair, Graeme, et al. 2019. *Estimatr: Fast Estimators for Design-Based Inference.* R package version 0.16.

Blanchflower, David G. 2004. "Self-Employment: More May Not Be Better". *Swedish Economic Policy Review* 11 (2): 15–73.

Blanchflower, David G., and Andrew J. Oswald. 1998. "What Makes an Entrepreneur?" *Journal of Labor Economics* 16 (1): 26–60. doi:10.1086/209881.

Block, Jörn, and Philipp Sandner. 2009. "Necessity and Opportunity Entrepreneurs and Their Duration in Self-Employment: Evidence from German Micro Data". *Journal of Industry, Competition and Trade* 9 (2): 117–137. doi:10.1007/s10842-007-0029-3.

Cahuc, Pierre, Stéphane Carcillo, and André Zylberberg. 2014. *Labor Economics.* 2nd edition. Cambridge, MA: MIT Press.

Chernozhukov, Victor. 2005. "Extremal Quantile Regression". *The Annals of Statistics* 33 (2): 806–839.

Cho, Yoon Y., David A. Robalino, and Jose M. Romero. 2015. *Entering and Leaving Self-Employment: A Panel Data Analysis for 12 Developing Countries.* IZA Discussion Papers 9358. Institute for the Study of Labor (IZA).

Diamond, Peter A. 1982. "Wage Determination and Efficiency in Search Equilibrium". *The Review of Economic Studies* 49 (2): 217–227. doi:10.2307/2297271.

Earle, John S, and Zuzana Sakova. 2000. "Business Start-Ups or Disguised Unemployment? Evidence on the Character of Self-Employment from Transition Economies". *Labour Economics* 7 (5): 575–601. doi:10.1016/S0927-5371(00)00014-2.

Ekelund, Jesper, et al. 2005. "Self-Employment and Risk Aversion—Evidence from Psychological Test Data". *Labour Economics* 12 (5): 649–659. doi:10.1016/j.labeco.2004.02.009.

Evans, David S., and Boyan Jovanovic. 1989. "An Estimated Model of Entrepreneurial Choice under Liquidity Constraints". *Journal of Political Economy* 97 (4): 808–827.

Evans, David S., and Linda S. Leighton. 1989. "Some Empirical Aspects of Entrepreneurship". *The American Economic Review* 79 (3): 519–535.

Fairlie, Robert W., and Frank M. Fossen. 2018. *Defining Opportunity Versus Necessity Entrepreneurship: Two Components of Business Creation.* Working Paper 17-014. Institute for Economic Policy Research (SIEPR).

Fields, Gary S. 2011. *Working Hard, Working Poor: A Global Journey.* Oxford University Press.

Gindling, T. H., and David Newhouse. 2014. "Self-Employment in the Developing World". *World Development* 56:313–331. doi:10.1016/j.worlddev.2013.03.003.

Global Entrepreneurship Monitor. 2017. *Empreendedorismo No Brasil: 2016.* Curitiba: IBQP.

Grimm, Michael, Peter Knorringa, and Jann Lay. 2012. *Constrained Gazelles : High Potentials in West Africa's Informal Economy* 77926. The World Bank.

Gronau, Reuben. 1971. "Information and Frictional Unemployment". *The American Economic Review* 61 (3): 290–301.

Hamilton, Barton H. 2000. "Does Entrepreneurship Pay? An Empirical Analysis of the Returns to Self-Employment". *Journal of Political Economy* 108 (3): 604–631.

Hanglberger, Dominik, and Joachim Merz. 2015. "Does Self-Employment Really Raise Job Satisfaction? Adaptation and Anticipation Effects on Self-Employment and General Job Changes". *Journal for Labour Market Research* 48 (4): 287–303. doi:10.1007/s12651-015-0175-8.

Hausman, Jerry A., et al. 2019. *Errors in the Dependent Variable of Quantile Regression Models.* Working Paper 25819. National Bureau of Economic Research.

Heckman, James J. 1979. "Sample Selection Bias as a Specification Error". *Econometrica* 47 (1): 153–161. doi:10.2307/1912352.

Hoffmann, Eivind. 2003. "International Classification of Status in Employment, ICSE-93". In *Advances in Cross-National Comparison: A European Working Book for Demographic and Socio-Economic Variables*, edited by Jürgen H. P. Hoffmeyer-Zlotnik and Christof Wolf, 125–136. Boston, MA: Springer US.

IBGE. 2016. *Pesquisa Nacional Por Amostra de Domicílios Contínua: Manual Básico Da Entrevista.* Rio de Janeiro. https://biblioteca.ibge.gov.br/visualizacao/instrumentos_de_coleta/doc5361.pdf.

— . 2014. *Pesquisa Nacional Por Amostra de Domicílios Contínua: Notas Metodológicas.* Rio de Janeiro. ftp://ftp.ibge.gov.br/Trabalho_e_Rendimento/Pesquisa_Nacional_por_Amostra_de_Domicilios_continua/Notas_metodologicas/notas_metodologicas.pdf.

— . 2018. *Pesquisa Nacional Por Amostra de Domicílios Contínua: Notas Técnicas v1.5.* Rio de Janeiro. https://biblioteca.ibge.gov.br/visualizacao/livros/liv101548_notas_tecnicas.pdf.

Kalbfleisch, J. D., and Ross L. Prentice. 2002. *The Statistical Analysis of Failure Time Data.* 2nd edition. Hoboken, N.J: J. Wiley.

Kanbur, S. M. Ravi. 1982. "Entrepreneurial Risk Taking, Inequality, and Public Policy: An Application of Inequality Decomposition Analysis to the General Equilibrium Effects of Progressive Taxation". *Journal of Political Economy* 90 (1): 1–21.

Kiefer, Nicholas M. 1988. "Economic Duration Data and Hazard Functions". *Journal of Economic Literature* 26 (2): 646–679.

Kihlstrom, Richard E., and Jean-Jacques Laffont. 1979. "A General Equilibrium Entrepreneurial Theory of Firm Formation Based on Risk Aversion". *Journal of Political Economy* 87 (4): 719–748. doi:10.1086/260790.

Knight, Frank H. 1921. *Risk, Uncertainty and Profit.* Boston & New York: Houghton Mifflin Co.

Koenker, Roger. 2018. *Quantreg: Quantile Regression.* R package version 5.38.

Lazear, Edward P. 2005. "Entrepreneurship". *Journal of Labor Economics* 23 (4): 649–680. doi:10.1086/491605.

— . 1979. "Why Is There Mandatory Retirement?" *Journal of Political Economy* 87 (6): 1261–1284.

Lindh, Thomas, and Henry Ohlsson. 1996. "Self-Employment and Windfall Gains: Evidence from the Swedish Lottery". *The Economic Journal* 106 (439): 1515–1526. doi:10.2307/2235198.

Lucas, Robert E. 1978. "On the Size Distribution of Business Firms". *The Bell Journal of Economics* 9 (2): 508–523. doi:10.2307/3003596.

Mandelman, Federico S., and Gabriel V. Montes-Rojas. 2009. "Is Self-Employment and Micro-Entrepreneurship a Desired Outcome?" *World Development* 37 (12): 1914–1925. doi:10.1016/j.worlddev.2009.05.005.

Margolis, David N. 2014. “By Choice and by Necessity: Entrepreneurship and Self-Employment in the Developing World”. *The European Journal of Development Research* 26 (4): 419–436. doi:10.1057/ejdr.2014.25.

— . 2009. *The Structure of Unemployment Compensation, Labor Market Signaling and Informal Sector Work.* Mimeo.

Margolis, David N., Lucas Navarro, and David A. Robalino. 2012. *Unemployment Insurance, Job Search and Informal Employment.* IZA Discussion Papers 6660. Institute for the Study of Labor (IZA).

Masclet, David, et al. 2009. “Group and Individual Risk Preferences: A Lottery-Choice Experiment with Self-Employed and Salaried Workers”. *Journal of Economic Behavior & Organization*, Field Experiments in Economics, 70 (3): 470–484. doi:10.1016/j.jebo.2007.11.002.

McCall, J. J. 1970. “Economics of Information and Job Search”. *The Quarterly Journal of Economics* 84 (1): 113–126. doi:10.2307/1879403.

Menezes-Filho, Naércio A., and Paulo Picchetti. 2000. “Os Determinantes Da Duração Do Desemprego Em São Paulo”. *Pesquisa e Planejamento Econômico* 30 (1): 23–48.

Meyer, Bruce D. 1990. “Unemployment Insurance and Unemployment Spells”. *Econometrica* 58 (4): 757–782. doi:10.2307/2938349.

Mincer, Jacob. 1974. *Schooling, Experience, and Earnings.* New York: Natonal Bureau of Economic Research.

Mondragón-Vélez, Camilo, and Ximena Peña. 2010. “Business Ownership and Self-Employment in Developing Economies: The Colombian Case”. In *International Differences in Entrepreneurship*, edited by Josh Lerner and Antoinette Schoar, 89–127. University of Chicago Press.

Monsen, Erik, Prashanth Mahagaonkar, and Christian Dienes. 2012. “Entrepreneurship in India: The Question of Occupational Transition”. *Small Business Economics* 39 (2): 359–382. doi:10.1007/s11187-011-9316-9.

Mortensen, Dale T. 1987. “Job Search and Labor Market Analysis”. In *Handbook of Labor Economics*, edited by Orley Ashenfelter and Richard Layard, 2:849–919. Amsterdam: North Holland.

— . 1982. “The Matching Process as a Noncooperative Bargaining Game”. In *The Economics of Information and Uncertainty*, edited by John J. McCall, 233–258. University of Chicago Press.

Mortensen, Dale T., and Christopher A. Pissarides. 1994. "Job Creation and Job Destruction in the Theory of Unemployment". *The Review of Economic Studies* 61 (3): 397–415. doi:10.2307/2297896.

Narita, Renata. 2013. *Self Employment in Developing Countries: A Search-Equilibrium Approach.* Working Paper n 2013-21. University of São Paulo (FEA-USP).

Pissarides, Christopher A. 2000. *Equilibrium Unemployment Theory.* 2nd edition. Cambridge, Mass: MIT Press.

— . 1984. "Search Intensity, Job Advertising, and Efficiency". *Journal of Labor Economics* 2 (1): 128–143.

Poschke, Markus. 2013. "'Entrepreneurs out of Necessity': A Snapshot". *Applied Economics Letters* 20 (7): 658–663. doi:10.1080/13504851.2012.727968.

Ramnath, Shanthi, John B Shoven, and Sita Nataraj Slavov. 2017. *Pathways to Retirement through Self-Employment.* Working Paper 23551. National Bureau of Economic Research.

Ribas, Rafael Perez, and Sergei Suarez Dillon Soares. 2008. *Sobre o Painel Da Pesquisa Mensal de Emprego (PME) Do IBGE.* Texto para Discussão TD 1348. Rio de Janeiro: IPEA.

Rogerson, Richard, Robert Shimer, and Randall Wright. 2005. "Search-Theoretic Models of the Labor Market: A Survey". *Journal of Economic Literature* 43 (4): 959–988.

Roy, A. D. 1951. "Some Thoughts on the Distribution of Earnings". *Oxford Economic Papers* 3 (2): 135–146.

Schoar, Antoinette. 2010. "The Divide between Subsistence and Transformational Entrepreneurship". *Innovation Policy and the Economy* 10:57–81.

StataCorp. 2015. *Stata Statistical Software.* Release 14. College Station, TX: StataCorp LP.

Stigler, George J. 1961. "The Economics of Information". *Journal of Political Economy* 69 (3): 213–225.

Terrell, Katherine, and Michael Troilo. 2010. "Values and Female Entrepreneurship". *International Journal of Gender and Entrepreneurship* 2 (3): 260–286. doi:10.1108/17566261011079242.

A Appendix

Table 3: *OLS regression of log monthly wage on individual attributes, using data from 2 samples: (a) all employees (baseline) and (b) employees with less than 12 months on the job*

	Sample: All employees		**Sample: New hires**	
term	estimate	std.error	estimate	std.error
(Intercept)	5.65***	(0.02)	5.79***	(0.02)
*Gender and age group (reference: female * 19 and below)*				
female * 20-24	0.23***	(0.01)	0.19***	(0.01)
female * 25-29	0.35***	(0.01)	0.28***	(0.01)
female * 30-34	0.43***	(0.01)	0.30***	(0.01)
female * 35-39	0.49***	(0.01)	0.30***	(0.01)
female * 40-44	0.54***	(0.01)	0.31***	(0.01)
female * 45-49	0.56***	(0.01)	0.30***	(0.01)
female * 50-54	0.60***	(0.01)	0.26***	(0.01)
female * 55-59	0.62***	(0.01)	0.20***	(0.02)
female * 60 and above	0.61***	(0.01)	0.11***	(0.02)
male * 19 and below	0.23***	(0.01)	0.21***	(0.01)
male * 20-24	0.50***	(0.01)	0.44***	(0.01)
male * 25-29	0.67***	(0.01)	0.57***	(0.01)
male * 30-34	0.80***	(0.01)	0.66***	(0.01)
male * 35-39	0.88***	(0.01)	0.71***	(0.01)
male * 40-44	0.95***	(0.01)	0.77***	(0.01)
male * 45-49	1.00***	(0.01)	0.74***	(0.01)
male * 50-54	1.03***	(0.01)	0.73***	(0.01)
male * 55-59	1.04***	(0.01)	0.68***	(0.01)
male * 60 and above	1.02***	(0.01)	0.63***	(0.02)
*Race and education (reference: non-white * no school)*				
non-white * elem school incomp	-0.01**	(0.01)	-0.04***	(0.01)
non-white * elem school	0.17***	(0.01)	0.12***	(0.01)
non-white * high school incomp	0.22***	(0.01)	0.16***	(0.01)
non-white * high school	0.39***	(0.01)	0.28***	(0.01)
non-white * college no degree	0.59***	(0.01)	0.42***	(0.01)
non-white * graduate or above	1.09***	(0.01)	0.86***	(0.01)
white * no school	0.09***	(0.01)	0.05***	(0.02)
white * elem school incomp	0.06***	(0.01)	0.03***	(0.01)
white * elem school	0.25***	(0.01)	0.19***	(0.01)
white * high school incomp	0.30***	(0.01)	0.22***	(0.01)
white * high school	0.47***	(0.01)	0.34***	(0.01)
white * college no degree	0.62***	(0.01)	0.40***	(0.01)
white * graduate or above	1.22***	(0.01)	0.95***	(0.01)

Table 3: *OLS regression of log monthly wage on individual attributes, using data from 2 samples: (a) all employees (baseline) and (b) employees with less than 12 months on the job (continued)*

	Sample: All employees		**Sample: New hires**	
term	estimate	std.error	estimate	std.error
Literacy (reference: illiterate)				
literate	0.37***	(0.01)	0.32***	(0.01)
At school (reference: not at school)				
private	-0.03***	(0.00)	-0.05***	(0.01)
public	-0.15***	(0.00)	-0.22***	(0.01)
Urban status (reference: rural)				
urban	0.15***	(0.00)	0.16***	(0.01)
*State area (reference: not capital * RO)*				
not capital * AC	-0.18***	(0.02)	-0.23***	(0.03)
not capital * AM	-0.23***	(0.02)	-0.21***	(0.03)
not capital * RR	-0.04*	(0.02)	0.01	(0.03)
not capital * PA	-0.13***	(0.02)	-0.19***	(0.02)
not capital * AP	0.01	(0.03)	-0.04	(0.03)
not capital * TO	-0.12***	(0.02)	-0.14***	(0.02)
not capital * MA	-0.35***	(0.02)	-0.45***	(0.02)
not capital * PI	-0.47***	(0.02)	-0.61***	(0.03)
not capital * CE	-0.43***	(0.02)	-0.50***	(0.02)
not capital * RN	-0.26***	(0.02)	-0.32***	(0.03)
not capital * PB	-0.36***	(0.02)	-0.44***	(0.03)
not capital * PE	-0.26***	(0.02)	-0.31***	(0.02)
not capital * AL	-0.23***	(0.02)	-0.30***	(0.02)
not capital * SE	-0.30***	(0.02)	-0.36***	(0.03)
not capital * BA	-0.31***	(0.02)	-0.38***	(0.02)
not capital * MG	-0.05***	(0.01)	-0.04**	(0.02)
not capital * ES	-0.00	(0.01)	0.02	(0.02)
not capital * RJ	0.04***	(0.01)	0.07***	(0.02)
not capital * SP	0.13***	(0.01)	0.14***	(0.02)
not capital * PR	0.10***	(0.01)	0.12***	(0.02)
not capital * SC	0.20***	(0.01)	0.26***	(0.02)
not capital * RS	0.12***	(0.01)	0.14***	(0.02)
not capital * MS	0.13***	(0.02)	0.16***	(0.02)
not capital * MT	0.18***	(0.02)	0.23***	(0.02)
not capital * GO	0.07***	(0.02)	0.11***	(0.02)
capital * RO	0.10***	(0.02)	0.09***	(0.02)
capital * AC	0.03	(0.02)	-0.03	(0.02)
capital * AM	-0.02	(0.02)	-0.08***	(0.02)
capital * RR	0.06**	(0.02)	-0.03	(0.02)
capital * PA	-0.08***	(0.02)	-0.15***	(0.02)

Table 3: *OLS regression of log monthly wage on individual attributes, using data from 2 samples: (a) all employees (baseline) and (b) employees with less than 12 months on the job (continued)*

	Sample: All employees		**Sample: New hires**	
term	estimate	std.error	estimate	std.error
capital * AP	0.10***	(0.03)	-0.09***	(0.03)
capital * TO	0.12***	(0.03)	0.09***	(0.02)
capital * MA	-0.08***	(0.02)	-0.06***	(0.02)
capital * PI	-0.06**	(0.02)	-0.09***	(0.02)
capital * CE	-0.06***	(0.02)	-0.08***	(0.02)
capital * RN	-0.09***	(0.02)	-0.14***	(0.02)
capital * PB	-0.02	(0.03)	-0.06**	(0.02)
capital * PE	0.02	(0.03)	-0.04	(0.03)
capital * AL	-0.08***	(0.02)	-0.07***	(0.02)
capital * SE	0.06*	(0.03)	-0.00	(0.03)
capital * BA	-0.06***	(0.02)	-0.11***	(0.02)
capital * MG	0.13***	(0.02)	0.08***	(0.02)
capital * ES	0.23***	(0.03)	0.14***	(0.03)
capital * RJ	0.21***	(0.02)	0.22***	(0.02)
capital * SP	0.23***	(0.02)	0.24***	(0.02)
capital * PR	0.24***	(0.02)	0.21***	(0.02)
capital * SC	0.26***	(0.02)	0.26***	(0.02)
capital * RS	0.26***	(0.02)	0.20***	(0.02)
capital * MS	0.14***	(0.02)	0.17***	(0.02)
capital * MT	0.14***	(0.02)	0.13***	(0.02)
capital * GO	0.14***	(0.02)	0.15***	(0.02)
capital * DF	0.35***	(0.02)	0.18***	(0.02)
Model statistics				
Observations	1,218,241		298,774	
Number of clusters	19,189		18,812	
R-Squared	0.49		0.37	
Adj. R-Squared	0.49		0.37	
F-statistic	1,581		632	
Degrees of freedom	88		88	
Model p-value	0.000		0.000	

Notes: Standard errors clustered at the level of the primary sampling unit of the survey. Statistical significance denoted by (*) for p < 0.1; (**) for p < 0.05; and (***) for p < 0.01. Estimated with the R package *estimatr* (Blair et al., 2019).

Table 4: *Quantile regression of log monthly wage on individual attributes for 5th, 10th (baseline) and 15th percentiles, using data from all employees*

	Quantile = 0.05		**Quantile = 0.10**		**Quantile = 0.15**	
term	estimate	std.error	estimate	std.error	estimate	std.error
(Intercept)	4.32***	(0.03)	4.69***	(0.02)	4.99***	(0.02)
Position in the household (reference: non relative)						
head with partner	0.12***	(0.02)	0.10***	(0.01)	0.08***	(0.01)
head without partner	-0.00	(0.02)	-0.00	(0.01)	-0.01	(0.01)
partner	0.06***	(0.02)	0.04***	(0.01)	0.03***	(0.01)
other relative	-0.04**	(0.02)	-0.05***	(0.01)	-0.06***	(0.01)
child	-0.11***	(0.01)	-0.11***	(0.01)	-0.12***	(0.01)
Composition of the household						
n. of adults(24 < age < 66)	0.02***	(0.00)	0.01***	(0.00)	0.01***	(0.00)
n. of young members (age < 25)	-0.03***	(0.00)	-0.02***	(0.00)	-0.02***	(0.00)
n. of senior members (age > 65)	-0.03***	(0.00)	-0.03***	(0.00)	-0.03***	(0.00)
*Gender and age group (reference: female * 19 and below)*						
female * 20-24	0.32***	(0.01)	0.34***	(0.01)	0.36***	(0.01)
female * 25-29	0.44***	(0.01)	0.46***	(0.01)	0.46***	(0.01)
female * 30-34	0.48***	(0.01)	0.51***	(0.01)	0.50***	(0.01)
female * 35-39	0.52***	(0.01)	0.53***	(0.01)	0.52***	(0.01)
female * 40-44	0.56***	(0.01)	0.57***	(0.01)	0.55***	(0.01)
female * 45-49	0.55***	(0.02)	0.56***	(0.01)	0.54***	(0.01)
female * 50-54	0.52***	(0.02)	0.56***	(0.01)	0.56***	(0.01)
female * 55-59	0.48***	(0.02)	0.52***	(0.01)	0.53***	(0.01)
female * 60 and above	0.32***	(0.02)	0.39***	(0.01)	0.45***	(0.01)
male * 19 and below	0.28***	(0.02)	0.27***	(0.01)	0.24***	(0.01)
male * 20-24	0.67***	(0.01)	0.68***	(0.01)	0.65***	(0.01)
male * 25-29	0.79***	(0.01)	0.74***	(0.01)	0.70***	(0.01)
male * 30-34	0.87***	(0.01)	0.81***	(0.01)	0.77***	(0.01)
male * 35-39	0.93***	(0.01)	0.87***	(0.01)	0.83***	(0.01)
male * 40-44	0.97***	(0.01)	0.92***	(0.01)	0.89***	(0.01)
male * 45-49	0.99***	(0.01)	0.94***	(0.01)	0.90***	(0.01)
male * 50-54	0.99***	(0.01)	0.94***	(0.01)	0.91***	(0.01)
male * 55-59	0.95***	(0.02)	0.92***	(0.01)	0.89***	(0.01)
male * 60 and above	0.85***	(0.02)	0.86***	(0.01)	0.85***	(0.01)
*Race and education (reference: non-white * no school)*						
non-white * elem school incomp	-0.05***	(0.02)	-0.06***	(0.01)	-0.05***	(0.01)
non-white * elem school	0.29***	(0.02)	0.23***	(0.01)	0.19***	(0.01)
non-white * high school incomp	0.34***	(0.02)	0.28***	(0.01)	0.24***	(0.01)
non-white * high school	0.57***	(0.02)	0.46***	(0.01)	0.39***	(0.01)
non-white * college no degree	0.71***	(0.02)	0.58***	(0.01)	0.51***	(0.01)

Table 4: *Quantile regression of log monthly wage on individual attributes for 5th, 10th (baseline) and 15th percentiles, using data from all employees (continued)*

	Quantile = 0.05		**Quantile = 0.10**		**Quantile = 0.15**	
term	estimate	std.error	estimate	std.error	estimate	std.error
non-white * graduate or above	1.03***	(0.02)	0.89***	(0.01)	0.82***	(0.01)
white * no school	0.12***	(0.02)	0.13***	(0.02)	0.10***	(0.01)
white * elem school incomp	0.06***	(0.02)	0.07***	(0.01)	0.06***	(0.01)
white * elem school	0.42***	(0.02)	0.32***	(0.01)	0.26***	(0.01)
white * high school incomp	0.46***	(0.02)	0.38***	(0.01)	0.32***	(0.01)
white * high school	0.64***	(0.02)	0.51***	(0.01)	0.42***	(0.01)
white * college no degree	0.68***	(0.02)	0.55***	(0.01)	0.48***	(0.01)
white * graduate or above	1.06***	(0.02)	0.93***	(0.01)	0.89***	(0.01)
Literacy (reference: illiterate)						
literate	0.51***	(0.02)	0.55***	(0.01)	0.56***	(0.01)
At school (reference: not at school)						
private	-0.03***	(0.01)	-0.03***	(0.00)	-0.03***	(0.00)
public	-0.25***	(0.01)	-0.24***	(0.01)	-0.22***	(0.01)
Urban status (reference: rural)						
urban	0.36***	(0.01)	0.29***	(0.00)	0.24***	(0.00)
*State area (reference: not capital * RO)*						
not capital * AC	-0.33***	(0.03)	-0.22***	(0.02)	-0.21***	(0.02)
not capital * AM	-0.30***	(0.03)	-0.22***	(0.02)	-0.22***	(0.01)
not capital * RR	-0.01	(0.02)	-0.04**	(0.02)	-0.04**	(0.02)
not capital * PA	-0.25***	(0.02)	-0.17***	(0.01)	-0.16***	(0.01)
not capital * AP	0.05*	(0.03)	0.04***	(0.02)	0.01	(0.02)
not capital * TO	-0.16***	(0.02)	-0.13***	(0.01)	-0.13***	(0.01)
not capital * MA	-0.63***	(0.02)	-0.54***	(0.01)	-0.50***	(0.01)
not capital * PI	-0.83***	(0.02)	-0.78***	(0.02)	-0.74***	(0.02)
not capital * CE	-0.77***	(0.02)	-0.64***	(0.01)	-0.58***	(0.01)
not capital * RN	-0.46***	(0.02)	-0.33***	(0.02)	-0.28***	(0.01)
not capital * PB	-0.63***	(0.02)	-0.54***	(0.02)	-0.47***	(0.01)
not capital * PE	-0.41***	(0.02)	-0.27***	(0.02)	-0.24***	(0.01)
not capital * AL	-0.33***	(0.02)	-0.22***	(0.02)	-0.22***	(0.01)
not capital * SE	-0.50***	(0.02)	-0.41***	(0.02)	-0.37***	(0.02)
not capital * BA	-0.57***	(0.02)	-0.47***	(0.01)	-0.42***	(0.01)
not capital * MG	-0.04**	(0.01)	0.00	(0.01)	0.00	(0.01)
not capital * ES	0.11***	(0.02)	0.11***	(0.01)	0.08***	(0.01)
not capital * RJ	0.18***	(0.01)	0.17***	(0.01)	0.12***	(0.01)
not capital * SP	0.23***	(0.01)	0.25***	(0.01)	0.20***	(0.01)
not capital * PR	0.19***	(0.02)	0.20***	(0.01)	0.17***	(0.01)
not capital * SC	0.37***	(0.01)	0.34***	(0.01)	0.29***	(0.01)
not capital * RS	0.23***	(0.02)	0.22***	(0.01)	0.18***	(0.01)

Table 4: *Quantile regression of log monthly wage on individual attributes for 5th, 10th (baseline) and 15th percentiles, using data from all employees (continued)*

	Quantile = 0.05		**Quantile = 0.10**		**Quantile = 0.15**	
term	estimate	std.error	estimate	std.error	estimate	std.error
not capital * MS	0.16***	(0.02)	0.18***	(0.01)	0.15***	(0.01)
not capital * MT	0.26***	(0.02)	0.25***	(0.01)	0.20***	(0.01)
not capital * GO	0.12***	(0.02)	0.13***	(0.01)	0.11***	(0.01)
capital * RO	0.18***	(0.02)	0.19***	(0.01)	0.14***	(0.01)
capital * AC	0.07***	(0.02)	0.06***	(0.01)	0.05***	(0.01)
capital * AM	0.02	(0.02)	0.02	(0.01)	0.01	(0.01)
capital * RR	0.08***	(0.02)	0.06***	(0.01)	0.04***	(0.01)
capital * PA	-0.09***	(0.02)	-0.06***	(0.02)	-0.07***	(0.01)
capital * AP	0.06***	(0.02)	0.06***	(0.02)	0.04***	(0.01)
capital * TO	0.22***	(0.02)	0.19***	(0.01)	0.13***	(0.01)
capital * MA	0.02	(0.02)	-0.00	(0.01)	-0.03**	(0.01)
capital * PI	-0.01	(0.02)	-0.02	(0.02)	-0.04***	(0.01)
capital * CE	0.01	(0.02)	-0.00	(0.01)	-0.03**	(0.01)
capital * RN	-0.03	(0.02)	-0.04***	(0.02)	-0.05***	(0.01)
capital * PB	0.02	(0.02)	-0.02	(0.02)	-0.02*	(0.01)
capital * PE	0.04*	(0.02)	0.04***	(0.01)	0.02*	(0.01)
capital * AL	0.03*	(0.02)	0.01	(0.01)	-0.03**	(0.01)
capital * SE	0.13***	(0.02)	0.10***	(0.02)	0.07***	(0.01)
capital * BA	-0.10***	(0.02)	-0.07***	(0.01)	-0.08***	(0.01)
capital * MG	0.22***	(0.02)	0.21***	(0.01)	0.16***	(0.01)
capital * ES	0.22***	(0.02)	0.23***	(0.02)	0.19***	(0.01)
capital * RJ	0.30***	(0.01)	0.27***	(0.01)	0.22***	(0.01)
capital * SP	0.30***	(0.02)	0.29***	(0.01)	0.25***	(0.01)
capital * PR	0.33***	(0.02)	0.31***	(0.01)	0.27***	(0.01)
capital * SC	0.35***	(0.02)	0.34***	(0.01)	0.30***	(0.01)
capital * RS	0.28***	(0.02)	0.27***	(0.01)	0.23***	(0.01)
capital * MS	0.26***	(0.02)	0.24***	(0.01)	0.18***	(0.01)
capital * MT	0.31***	(0.02)	0.25***	(0.01)	0.19***	(0.01)
capital * GO	0.23***	(0.02)	0.22***	(0.01)	0.17***	(0.01)
capital * DF	0.31***	(0.02)	0.29***	(0.01)	0.25***	(0.01)
Model statistics						
Observations	1,218,241		1,218,241		1,218,241	
Log-likelihood	-1,621,653		-1,406,317		-1,275,172	
AIC	3,243,501		2,812,829		2,550,538	
BIC	3,244,666		2,813,995		2,551,703	
Degrees of freedom	97		97		97	

Notes: Statistical significance denoted by (*) for p < 0.1; (**) for p < 0.05; and (***) for p < 0.01. Estimated by the Frisch-Newton interior point method after preprocessing, with the R package *quantreg* (Koenker, 2018).

Table 5: *Quantile regression of log monthly wage on individual attributes for 5th, 10th and 15th percentiles, using data from employees with less than 12 months on the job*

	Quantile = 0.05		Quantile = 0.10		Quantile = 0.15	
term	estimate	std.error	estimate	std.error	estimate	std.error
(Intercept)	4.50***	(0.06)	4.88***	(0.04)	5.13***	(0.03)
Position in the household (reference: non relative)						
head with partner	0.09**	(0.04)	0.05**	(0.02)	0.03*	(0.02)
head without partner	-0.04	(0.04)	-0.05**	(0.02)	-0.05***	(0.02)
partner	0.04	(0.04)	0.01	(0.02)	-0.00	(0.02)
other relative	-0.04	(0.04)	-0.06***	(0.02)	-0.08***	(0.02)
child	-0.13***	(0.04)	-0.13***	(0.02)	-0.14***	(0.02)
Composition of the household						
n. of adults(24 < age < 66)	0.01**	(0.01)	0.01***	(0.00)	0.01***	(0.00)
n. of young members (age < 25)	-0.02***	(0.00)	-0.02***	(0.00)	-0.02***	(0.00)
n. of senior members (age > 65)	-0.03***	(0.01)	-0.03***	(0.01)	-0.03***	(0.01)
*Gender and age group (reference: female * 19 and below)*						
female * 20-24	0.17***	(0.02)	0.23***	(0.01)	0.25***	(0.01)
female * 25-29	0.21***	(0.02)	0.26***	(0.02)	0.30***	(0.01)
female * 30-34	0.18***	(0.03)	0.26***	(0.02)	0.30***	(0.01)
female * 35-39	0.20***	(0.03)	0.25***	(0.02)	0.27***	(0.01)
female * 40-44	0.14***	(0.03)	0.20***	(0.02)	0.24***	(0.02)
female * 45-49	0.12***	(0.03)	0.19***	(0.02)	0.22***	(0.02)
female * 50-54	-0.01	(0.03)	0.09***	(0.03)	0.11***	(0.02)
female * 55-59	-0.13***	(0.04)	-0.09***	(0.03)	-0.05	(0.03)
female * 60 and above	-0.40***	(0.06)	-0.29***	(0.05)	-0.21***	(0.03)
male * 19 and below	0.26***	(0.02)	0.26***	(0.01)	0.25***	(0.01)
male * 20-24	0.52***	(0.02)	0.57***	(0.01)	0.55***	(0.01)
male * 25-29	0.62***	(0.02)	0.63***	(0.01)	0.60***	(0.01)
male * 30-34	0.69***	(0.02)	0.67***	(0.01)	0.64***	(0.01)
male * 35-39	0.74***	(0.02)	0.71***	(0.01)	0.68***	(0.01)
male * 40-44	0.77***	(0.02)	0.75***	(0.02)	0.72***	(0.01)
male * 45-49	0.72***	(0.03)	0.71***	(0.02)	0.68***	(0.01)
male * 50-54	0.63***	(0.03)	0.67***	(0.02)	0.65***	(0.02)
male * 55-59	0.49***	(0.06)	0.57***	(0.03)	0.56***	(0.02)
male * 60 and above	0.34***	(0.05)	0.42***	(0.03)	0.47***	(0.03)
*Race and education (reference: non-white * no school)*						
non-white * elem school incomp	-0.04	(0.03)	-0.08***	(0.02)	-0.09***	(0.02)
non-white * elem school	0.22***	(0.03)	0.18***	(0.03)	0.16***	(0.02)
non-white * high school incomp	0.31***	(0.03)	0.25***	(0.03)	0.21***	(0.02)
non-white * high school	0.50***	(0.03)	0.40***	(0.02)	0.35***	(0.02)
non-white * college no degree	0.65***	(0.03)	0.49***	(0.03)	0.42***	(0.02)

Table 5: *Quantile regression of log monthly wage on individual attributes for 5th, 10th and 15th percentiles, using data from employees with less than 12 months on the job (continued)*

	Quantile = 0.05		**Quantile = 0.10**		**Quantile = 0.15**	
term	estimate	std.error	estimate	std.error	estimate	std.error
non-white * graduate or above	0.98***	(0.03)	0.81***	(0.03)	0.72***	(0.02)
white * no school	0.10**	(0.04)	0.08**	(0.03)	0.07**	(0.03)
white * elem school incomp	0.01	(0.03)	-0.00	(0.03)	0.02	(0.02)
white * elem school	0.37***	(0.03)	0.29***	(0.03)	0.23***	(0.02)
white * high school incomp	0.42***	(0.03)	0.33***	(0.03)	0.28***	(0.02)
white * high school	0.59***	(0.03)	0.45***	(0.02)	0.39***	(0.02)
white * college no degree	0.61***	(0.03)	0.44***	(0.03)	0.38***	(0.02)
white * graduate or above	0.98***	(0.03)	0.82***	(0.02)	0.73***	(0.02)
Literacy (reference: illiterate)						
literate	0.36***	(0.03)	0.41***	(0.02)	0.44***	(0.02)
At school (reference: not at school)						
private	-0.06***	(0.01)	-0.09***	(0.01)	-0.09***	(0.01)
public	-0.27***	(0.02)	-0.29***	(0.01)	-0.30***	(0.01)
Urban status (reference: rural)						
urban	0.35***	(0.01)	0.32***	(0.01)	0.28***	(0.01)
*State area (reference: not capital * RO)*						
not capital * AC	-0.37***	(0.06)	-0.38***	(0.06)	-0.33***	(0.04)
not capital * AM	-0.35***	(0.05)	-0.31***	(0.04)	-0.26***	(0.04)
not capital * RR	0.04	(0.06)	0.09*	(0.05)	0.09***	(0.03)
not capital * PA	-0.44***	(0.05)	-0.35***	(0.03)	-0.29***	(0.03)
not capital * AP	0.05	(0.08)	0.10*	(0.05)	0.05*	(0.03)
not capital * TO	-0.16***	(0.05)	-0.17***	(0.03)	-0.17***	(0.03)
not capital * MA	-0.78***	(0.04)	-0.73***	(0.03)	-0.68***	(0.02)
not capital * PI	-0.88***	(0.05)	-0.90***	(0.03)	-0.88***	(0.03)
not capital * CE	-0.87***	(0.04)	-0.80***	(0.03)	-0.71***	(0.03)
not capital * RN	-0.71***	(0.06)	-0.56***	(0.04)	-0.48***	(0.04)
not capital * PB	-0.74***	(0.05)	-0.74***	(0.03)	-0.69***	(0.03)
not capital * PE	-0.59***	(0.05)	-0.50***	(0.03)	-0.41***	(0.03)
not capital * AL	-0.56***	(0.05)	-0.48***	(0.03)	-0.41***	(0.03)
not capital * SE	-0.63***	(0.05)	-0.55***	(0.04)	-0.52***	(0.03)
not capital * BA	-0.66***	(0.04)	-0.60***	(0.03)	-0.56***	(0.03)
not capital * MG	-0.09**	(0.04)	-0.04	(0.02)	0.01	(0.02)
not capital * ES	0.16***	(0.04)	0.15***	(0.02)	0.14***	(0.02)
not capital * RJ	0.27***	(0.04)	0.23***	(0.02)	0.20***	(0.02)
not capital * SP	0.23***	(0.04)	0.25***	(0.02)	0.25***	(0.02)
not capital * PR	0.22***	(0.04)	0.22***	(0.02)	0.23***	(0.02)
not capital * SC	0.51***	(0.04)	0.43***	(0.02)	0.40***	(0.02)
not capital * RS	0.30***	(0.04)	0.29***	(0.02)	0.27***	(0.02)

Table 5: *Quantile regression of log monthly wage on individual attributes for 5th, 10th and 15th percentiles, using data from employees with less than 12 months on the job (continued)*

	Quantile = 0.05		**Quantile = 0.10**		**Quantile = 0.15**	
term	estimate	std.error	estimate	std.error	estimate	std.error
not capital * MS	0.22***	(0.05)	0.23***	(0.03)	0.24***	(0.02)
not capital * MT	0.34***	(0.04)	0.33***	(0.02)	0.30***	(0.02)
not capital * GO	0.21***	(0.04)	0.20***	(0.02)	0.18***	(0.02)
capital * RO	0.16***	(0.06)	0.20***	(0.03)	0.20***	(0.03)
capital * AC	0.04	(0.05)	0.06*	(0.03)	0.05*	(0.03)
capital * AM	0.02	(0.05)	0.03	(0.03)	0.02	(0.02)
capital * RR	0.09*	(0.05)	0.12***	(0.03)	0.09***	(0.03)
capital * PA	-0.15***	(0.06)	-0.12***	(0.04)	-0.10***	(0.03)
capital * AP	-0.02	(0.08)	-0.07	(0.05)	-0.01	(0.03)
capital * TO	0.35***	(0.05)	0.27***	(0.03)	0.23***	(0.03)
capital * MA	0.13**	(0.05)	0.09***	(0.03)	0.05*	(0.02)
capital * PI	-0.04	(0.06)	-0.03	(0.03)	-0.04	(0.03)
capital * CE	0.06	(0.05)	0.04	(0.03)	0.03	(0.02)
capital * RN	-0.04	(0.06)	-0.10**	(0.04)	-0.06*	(0.03)
capital * PB	0.02	(0.07)	0.02	(0.03)	0.01	(0.03)
capital * PE	-0.04	(0.06)	-0.01	(0.04)	0.00	(0.03)
capital * AL	0.08	(0.05)	0.07*	(0.03)	0.06**	(0.02)
capital * SE	0.24***	(0.05)	0.16***	(0.03)	0.12***	(0.03)
capital * BA	-0.11*	(0.06)	-0.13***	(0.04)	-0.08***	(0.03)
capital * MG	0.22***	(0.05)	0.21***	(0.03)	0.20***	(0.02)
capital * ES	0.30***	(0.06)	0.22***	(0.03)	0.20***	(0.03)
capital * RJ	0.48***	(0.04)	0.36***	(0.02)	0.31***	(0.02)
capital * SP	0.37***	(0.05)	0.34***	(0.03)	0.32***	(0.02)
capital * PR	0.43***	(0.05)	0.35***	(0.03)	0.31***	(0.02)
capital * SC	0.43***	(0.05)	0.39***	(0.04)	0.38***	(0.03)
capital * RS	0.35***	(0.05)	0.29***	(0.03)	0.27***	(0.03)
capital * MS	0.41***	(0.05)	0.34***	(0.03)	0.31***	(0.02)
capital * MT	0.43***	(0.05)	0.34***	(0.03)	0.28***	(0.02)
capital * GO	0.41***	(0.04)	0.31***	(0.03)	0.28***	(0.03)
capital * DF	0.33***	(0.04)	0.27***	(0.03)	0.25***	(0.02)
Model statistics						
Observations	298,774		298,774		298,774	
Log-likelihood	-442,133		-386,867		-351,969	
AIC	884,461		773,929		704,133	
BIC	885,490		774,957		705,162	
Degrees of freedom	97		97		97	

Notes: Statistical significance denoted by (*) for p < 0.1; (**) for p < 0.05; and (***) for p < 0.01. Estimated by the Frisch-Newton interior point method after preprocessing, with the R package *quantreg* (Koenker, 2018).

Table 6: *Exponential duration models for unemployment and employment (baseline)*

	Unemployment Duration		**Employment Duration**	
	hazard ratio	std.error	hazard ratio	std.error
(Intercept)	0.37***	(0.039)	0.27***	(0.017)
Position in the household (reference: non relative)				
head with partner	0.99	(0.071)	0.94	(0.039)
head without partner	0.97	(0.070)	1.01	(0.042)
partner	0.94	(0.067)	1.07	(0.044)
other relative	0.84*	(0.061)	1.14**	(0.048)
child	0.74***	(0.052)	1.24***	(0.051)
Composition of the household				
n. of adults (24 < age < 66)	0.96***	(0.008)	1.01**	(0.005)
n. of young members (age < 25)	1.02***	(0.004)	1.04***	(0.003)
n. of senior members (age > 65)	0.93***	(0.015)	1.08***	(0.010)
*Gender and age group (reference: female * 19 and below)*				
female * 20-24	0.99	(0.029)	0.63***	(0.014)
female * 25-29	1.07*	(0.036)	0.53***	(0.012)
female * 30-34	1.10**	(0.038)	0.44***	(0.011)
female * 35-39	1.11**	(0.040)	0.39***	(0.010)
female * 40-44	1.12**	(0.043)	0.36***	(0.009)
female * 45-49	1.13**	(0.050)	0.36***	(0.009)
female * 50-54	1.17**	(0.059)	0.38***	(0.010)
female * 55-59	1.03	(0.068)	0.43***	(0.013)
female * 60 and above	1.14	(0.112)	0.60***	(0.019)
male * 19 and below	1.29***	(0.036)	0.83***	(0.017)
male * 20-24	1.27***	(0.035)	0.54***	(0.011)
male * 25-29	1.52***	(0.047)	0.48***	(0.011)
male * 30-34	1.50***	(0.050)	0.43***	(0.010)
male * 35-39	1.39***	(0.049)	0.41***	(0.010)
male * 40-44	1.30***	(0.049)	0.40***	(0.010)
male * 45-49	1.17***	(0.047)	0.39***	(0.010)
male * 50-54	1.07	(0.047)	0.41***	(0.011)
male * 55-59	1.01	(0.050)	0.45***	(0.013)
male * 60 and above	0.85*	(0.055)	0.58***	(0.016)
*Race and education (reference: non-white * no school)*				
non-white * elem school incomp	0.87***	(0.030)	0.87***	(0.017)
non-white * elem school	0.83***	(0.032)	0.74***	(0.017)
non-white * high school incomp	0.76***	(0.030)	0.65***	(0.016)
non-white * high school	0.73***	(0.026)	0.48***	(0.010)
non-white * college no degree	0.74***	(0.035)	0.41***	(0.012)
non-white * graduate or above	0.69***	(0.032)	0.26***	(0.007)
white * no school	0.93	(0.056)	1.02	(0.030)

Table 6: *Exponential duration models for unemployment and employment (baseline) (continued)*

	Unemployment Duration		**Employment Duration**	
	hazard ratio	std.error	hazard ratio	std.error
white * elem school incomp	0.86***	(0.034)	0.85***	(0.019)
white * elem school	0.85***	(0.038)	0.73***	(0.019)
white * high school incomp	0.78***	(0.035)	0.62***	(0.017)
white * high school	0.73***	(0.028)	0.49***	(0.011)
white * college no degree	0.78***	(0.039)	0.47***	(0.014)
white * graduate or above	0.66***	(0.030)	0.33***	(0.008)
Literacy (reference: illiterate)				
literate	0.86***	(0.029)	0.80***	(0.015)
At school (reference: not at school)				
private	1.10***	(0.031)	0.95**	(0.017)
public	0.95*	(0.021)	1.19***	(0.017)
Urban status (reference: rural)				
urban	0.86***	(0.015)	0.86***	(0.010)
*State area (reference: not capital * RO)*				
not capital * AC	0.73**	(0.081)	1.29***	(0.073)
not capital * AM	0.68***	(0.074)	1.40***	(0.078)
not capital * RR	0.92	(0.117)	1.31***	(0.100)
not capital * PA	0.83*	(0.064)	1.07	(0.054)
not capital * AP	0.29***	(0.048)	1.18	(0.114)
not capital * TO	0.95	(0.078)	1.06	(0.058)
not capital * MA	0.84*	(0.062)	1.43***	(0.065)
not capital * PI	1.00	(0.078)	1.38***	(0.072)
not capital * CE	0.79**	(0.058)	1.22***	(0.057)
not capital * RN	0.69***	(0.056)	0.97	(0.050)
not capital * PB	0.72***	(0.056)	0.94	(0.048)
not capital * PE	0.45***	(0.036)	1.04	(0.051)
not capital * AL	0.63***	(0.050)	1.19***	(0.062)
not capital * SE	0.59***	(0.056)	1.04	(0.056)
not capital * BA	0.59***	(0.045)	1.13**	(0.054)
not capital * MG	0.84*	(0.060)	0.75***	(0.034)
not capital * ES	0.65***	(0.050)	0.94	(0.045)
not capital * RJ	0.36***	(0.027)	0.70***	(0.033)
not capital * SP	0.57***	(0.042)	0.63***	(0.029)
not capital * PR	0.88	(0.066)	0.72***	(0.034)
not capital * SC	0.89	(0.066)	0.68***	(0.032)
not capital * RS	0.80**	(0.060)	0.64***	(0.030)
not capital * MS	1.08	(0.090)	0.62***	(0.036)
not capital * MT	0.98	(0.076)	0.82***	(0.042)
not capital * GO	0.83*	(0.063)	0.79***	(0.039)

Table 6: *Exponential duration models for unemployment and employment (baseline) (continued)*

	Unemployment Duration		**Employment Duration**	
	hazard ratio	std.error	hazard ratio	std.error
capital * RO	0.81*	(0.081)	1.07	(0.066)
capital * AC	0.74***	(0.063)	1.06	(0.056)
capital * AM	0.47***	(0.039)	1.13*	(0.061)
capital * RR	0.97	(0.101)	1.03	(0.060)
capital * PA	0.80*	(0.075)	1.04	(0.058)
capital * AP	0.40***	(0.062)	0.82**	(0.061)
capital * TO	0.77*	(0.090)	0.86*	(0.058)
capital * MA	0.54***	(0.048)	1.18**	(0.065)
capital * PI	1.09	(0.103)	0.84**	(0.049)
capital * CE	0.74***	(0.061)	0.97	(0.053)
capital * RN	0.52***	(0.055)	0.70***	(0.046)
capital * PB	0.66***	(0.067)	0.82**	(0.053)
capital * PE	0.61***	(0.062)	0.97	(0.055)
capital * AL	0.70***	(0.061)	1.58***	(0.092)
capital * SE	0.50***	(0.055)	1.15	(0.090)
capital * BA	0.69***	(0.060)	1.07	(0.060)
capital * MG	0.80*	(0.070)	0.86**	(0.046)
capital * ES	0.56***	(0.068)	0.97	(0.069)
capital * RJ	0.42***	(0.034)	0.72***	(0.038)
capital * SP	0.55***	(0.044)	0.73***	(0.039)
capital * PR	0.68***	(0.066)	0.93	(0.051)
capital * SC	0.86	(0.103)	1.11	(0.081)
capital * RS	0.91	(0.080)	0.87*	(0.048)
capital * MS	1.18	(0.111)	0.90	(0.055)
capital * MT	0.86	(0.089)	1.21**	(0.081)
capital * GO	1.01	(0.093)	0.95	(0.051)
capital * DF	0.53***	(0.045)	0.75***	(0.039)
Model statistics				
Number of individuals	67,190		373,172	
Number of transitions	35,594		113,538	
Number of clusters	15,449		18,094	
Chi square statistic	4,687		28,109	
Degrees of freedom	96		96	
Model p-value	0.000		0.000	

Notes: Standard errors clustered at the level of the primary sampling unit of the survey. Statistical significance denoted by (*) for p < 0.1; (**) for p < 0.05; and (***) for p < 0.01. Estimated using Stata (StataCorp, 2015).

Table 7: *Weibull duration models for unemployment and employment*

	Unemployment Duration		Employment Duration	
	hazard ratio	std.error	hazard ratio	std.error
(Intercept)	0.36***	(0.039)	0.41***	(0.026)
Position in the household (reference: non relative)				
head with partner	0.99	(0.071)	0.96	(0.039)
head without partner	0.97	(0.070)	1.02	(0.042)
partner	0.94	(0.067)	1.09*	(0.044)
other relative	0.84*	(0.061)	1.13**	(0.046)
child	0.74***	(0.052)	1.22***	(0.049)
Composition of the household				
n. of adults (24 < age < 66)	0.96***	(0.008)	1.03***	(0.005)
n. of young members (age < 25)	1.02***	(0.005)	1.03***	(0.003)
n. of senior members (age > 65)	0.93***	(0.015)	1.09***	(0.010)
*Gender and age group (reference: female * 19 and below)*				
female * 20-24	0.99	(0.029)	0.68***	(0.014)
female * 25-29	1.07*	(0.036)	0.59***	(0.014)
female * 30-34	1.10**	(0.038)	0.52***	(0.012)
female * 35-39	1.11**	(0.040)	0.48***	(0.011)
female * 40-44	1.12**	(0.043)	0.46***	(0.011)
female * 45-49	1.13**	(0.050)	0.48***	(0.012)
female * 50-54	1.17**	(0.059)	0.52***	(0.014)
female * 55-59	1.03	(0.068)	0.61***	(0.018)
female * 60 and above	1.13	(0.112)	0.90***	(0.027)
male * 19 and below	1.29***	(0.036)	0.86***	(0.017)
male * 20-24	1.27***	(0.035)	0.59***	(0.012)
male * 25-29	1.52***	(0.047)	0.55***	(0.012)
male * 30-34	1.50***	(0.050)	0.52***	(0.012)
male * 35-39	1.39***	(0.049)	0.51***	(0.012)
male * 40-44	1.30***	(0.050)	0.51***	(0.012)
male * 45-49	1.17***	(0.047)	0.52***	(0.013)
male * 50-54	1.07	(0.047)	0.57***	(0.015)
male * 55-59	1.01	(0.050)	0.64***	(0.018)
male * 60 and above	0.85*	(0.055)	0.85***	(0.024)
*Race and education (reference: non-white * no school)*				
non-white * elem school incomp	0.87***	(0.030)	0.87***	(0.017)
non-white * elem school	0.83***	(0.032)	0.76***	(0.017)
non-white * high school incomp	0.76***	(0.030)	0.67***	(0.016)
non-white * high school	0.73***	(0.026)	0.51***	(0.011)
non-white * college no degree	0.74***	(0.035)	0.45***	(0.013)
non-white * graduate or above	0.69***	(0.032)	0.29***	(0.008)
white * no school	0.93	(0.056)	1.03	(0.029)

Table 7: *Weibull duration models for unemployment and employment (continued)*

	Unemployment Duration		Employment Duration	
	hazard ratio	std.error	hazard ratio	std.error
white * elem school incomp	0.86***	(0.034)	0.86***	(0.019)
white * elem school	0.85***	(0.038)	0.76***	(0.019)
white * high school incomp	0.78***	(0.035)	0.65***	(0.017)
white * high school	0.73***	(0.028)	0.53***	(0.012)
white * college no degree	0.78***	(0.039)	0.51***	(0.015)
white * graduate or above	0.66***	(0.030)	0.37***	(0.009)
Literacy (reference: illiterate)				
literate	0.86***	(0.029)	0.83***	(0.015)
At school (reference: not at school)				
private	1.10***	(0.031)	0.95**	(0.017)
public	0.95*	(0.021)	1.18***	(0.017)
Urban status (reference: rural)				
urban	0.86***	(0.015)	0.86***	(0.009)
*State area (reference: not capital * RO)*				
not capital * AC	0.73**	(0.081)	1.31***	(0.073)
not capital * AM	0.68***	(0.074)	1.41***	(0.079)
not capital * RR	0.92	(0.117)	1.25**	(0.094)
not capital * PA	0.83*	(0.065)	1.05	(0.051)
not capital * AP	0.29***	(0.048)	1.17	(0.111)
not capital * TO	0.95	(0.079)	1.02	(0.055)
not capital * MA	0.84*	(0.062)	1.37***	(0.062)
not capital * PI	1.00	(0.079)	1.30***	(0.066)
not capital * CE	0.79**	(0.058)	1.21***	(0.055)
not capital * RN	0.69***	(0.056)	0.95	(0.048)
not capital * PB	0.72***	(0.056)	0.95	(0.047)
not capital * PE	0.45***	(0.036)	1.03	(0.050)
not capital * AL	0.62***	(0.050)	1.21***	(0.061)
not capital * SE	0.59***	(0.056)	1.07	(0.056)
not capital * BA	0.59***	(0.045)	1.12*	(0.052)
not capital * MG	0.84*	(0.061)	0.73***	(0.033)
not capital * ES	0.65***	(0.050)	0.93	(0.043)
not capital * RJ	0.36***	(0.027)	0.70***	(0.033)
not capital * SP	0.57***	(0.042)	0.63***	(0.029)
not capital * PR	0.88	(0.066)	0.72***	(0.033)
not capital * SC	0.89	(0.066)	0.68***	(0.031)
not capital * RS	0.80**	(0.060)	0.64***	(0.030)
not capital * MS	1.08	(0.091)	0.59***	(0.034)
not capital * MT	0.98	(0.076)	0.79***	(0.040)
not capital * GO	0.83*	(0.063)	0.78***	(0.038)

Table 7: *Weibull duration models for unemployment and employment (continued)*

	Unemployment Duration		Employment Duration	
	hazard ratio	std.error	hazard ratio	std.error
capital * RO	0.81*	(0.081)	1.07	(0.065)
capital * AC	0.74***	(0.064)	1.06	(0.056)
capital * AM	0.46***	(0.039)	1.10	(0.058)
capital * RR	0.97	(0.101)	1.01	(0.058)
capital * PA	0.81*	(0.075)	1.02	(0.056)
capital * AP	0.40***	(0.062)	0.83*	(0.062)
capital * TO	0.77*	(0.091)	0.84**	(0.056)
capital * MA	0.54***	(0.048)	1.16**	(0.063)
capital * PI	1.10	(0.104)	0.81***	(0.047)
capital * CE	0.74***	(0.061)	0.95	(0.051)
capital * RN	0.52***	(0.055)	0.71***	(0.045)
capital * PB	0.66***	(0.067)	0.83**	(0.053)
capital * PE	0.61***	(0.062)	0.98	(0.055)
capital * AL	0.70***	(0.061)	1.59***	(0.091)
capital * SE	0.50***	(0.055)	1.16	(0.091)
capital * BA	0.69***	(0.060)	1.03	(0.056)
capital * MG	0.81*	(0.070)	0.84***	(0.044)
capital * ES	0.55***	(0.068)	0.96	(0.068)
capital * RJ	0.42***	(0.035)	0.72***	(0.038)
capital * SP	0.55***	(0.044)	0.72***	(0.038)
capital * PR	0.68***	(0.067)	0.92	(0.050)
capital * SC	0.86	(0.104)	1.11	(0.079)
capital * RS	0.91	(0.080)	0.85**	(0.046)
capital * MS	1.19	(0.112)	0.88*	(0.052)
capital * MT	0.86	(0.090)	1.17*	(0.077)
capital * GO	1.01	(0.093)	0.94	(0.049)
capital * DF	0.53***	(0.045)	0.75***	(0.038)
Model ancillary parameter				
Weibull parameter	1.01	(0.004)	0.84***	(0.002)
Model statistics				
Number of individuals	67,190		373,172	
Number of transitions	35,594		113,538	
Number of clusters	15,449		18,094	
Chi square statistic	4,412		22,257	
Degrees of freedom	96		96	
Model p-value	0.000		0.000	

Notes: Standard errors clustered at the level of the primary sampling unit of the survey. Statistical significance denoted by (*) for p < 0.1; (**) for p < 0.05; and (***) for p < 0.01. Estimated using Stata (StataCorp, 2015).